PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/56, 15/11, 15/10, 15/64, 1/21, 15/63 — A1

(11) International Publication Number: **WO 98/06859**

(43) International Publication Date: 19 February 1998 (19.02.98)

(21) International Application Number: PCT/US96/13003

(22) International Filing Date: 9 August 1996 (09.08.96)

(71) Applicant *(for all designated States except US)*: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: THOTAKURA, Nageswara, Rao [US/US]; 7310 Turfside Lane, Gaithersburg, MD 20879 (US). CHOPRA, Arvind [CA/US]; 18 Ramsdell Court, Gaithersburg, MD 20878 (US). CHOI, Gil, H. [KR/US]; 11429 Potomac Oaks Drive, Rockville, MD 20850 (US). GENZT, Reiner, L. [DE/US]; 13404 Fairland Park Drive, Silver Spring, MD 20904 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US).

(74) Agents: GOLDSTEIN, Jorge, A. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).

(81) Designated States: AM, AU, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IL, JP, KG, KP, KR, KZ, LT, LV, MD, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, TJ, TM, TR, UA, US, UZ, VN, Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published

With international search report.

With an indication in relation to a deposited microorganism furnished under Rule 13bis separately from the description.

Date of receipt by the International Bureau:

29 August 1996 (29.08.96)

(54) Title: HUMAN CHITINASE ALPHA AND CHITINASE ALPHA-2

(57) Abstract

The present invention concerns novel human chitinase genes - chitinase alpha and chitinase alpha-2. More particularly, the invention concerns nucleic acid sequences encoding chitinase alpha, chitinase alpha-2, or fragments thereof, the use of such polypeptides for treating disease, as well as processes for producing these polypeptides recombinantly.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Human Chitinase Alpha and Chitinase Alpha-2

Background of the Invention

Field of the Invention

The present invention relates to novel human chitinase genes - chitinase alpha and chitinase alpha-2. More particularly, the invention relates to nucleic acid sequences encoding mature chitinase alpha or mature chitinase alpha-2 or fragments thereof, the use of such polypeptides for treating disease, as well as processes for producing these polypeptides recombinantly.

Related Art

The generation or destruction of tissue requires constant reorganization and restructuring of the extracellular matrix (ECM) components, including interstitial collagens, basement membrane collagen, fibronectin, laminin, aggrecan, and various proteoglycans. Heinegard and Oldberg, *FASEB J. 3*:2042-2051 (1989). Normal types of remodeling processes include embryonic development, post-partum involution of the uterus, ovulation, wound healing, and bone and growth plate remodeling. Woessner *et al.*, *Steroids 54*: 491-499 (1989); Weeks *et al.*, *Biochim. Biophys. Acta 445*:205-214 (1976); Lepage and Gache, *EMBO J.* 9:3003-3012 (1990); Wride and Sanders, *Dev-Dyn. 198*(3):225-239 (1993). Similar processes also occur in disease states such as joint destruction in rheumatoid and osteoarthritis, periodontia, and tumor cell metastasis. Thompson and Oegema, *J. Bone Joint Surg. 61*:407-416 (1979); Reynolds *et al.*, *Adv-Dent-Res. 8*(2):312-319 (1994). One example of these processes is the migration of macrophages to the site of inflammation as in the case of synovial tissue in rheumatoid arthritis. Cutolo *et al.*, *Clin. and Exper. Rheum.11*:331-339 (1993). The ECM components are regulated, in both normal and disease states, by various exogenous and endogenous factors. For example, in tumor formation, the

differentiation state of the cell can increase the rate of degradation of the ECM. Benya, *Pathol. Immunopathol. Res.* 7:51-54 (1988). Likewise, the presence of metalloproteinases or their inhibitors can alter the composition of the ECM. An imbalance of metalloproteinases and tissue inhibitors of matrix metalloproteinases (TIMP) has been shown to contribute to the pathogenesis of osteoarthritis. Dean *et al.*, *J. Clin Invest.* *84*:678-685 (1989). Cytokines, growth factors, and the extracellular environment can all contribute to the alteration of the ECM. Tyler, *Biochem J.* *227*:869-878 (1985); Dinarello, *Sem. Immunol.* *4*:133-145 (1992); McConnell *et al.*, *J. Cell Biol.* *105*:1087-1098-(1987).

The growth of cartilage and bone is actualized by cells such as articular chondrocytes and osteoblasts. The main function of these cells in immature tissue is the deposition and remodeling of the cartilage or bone matrix. In adult tissue, these cells maintain this matrix in order to ensure its proper function. In both case, this encompasses secretion of proteins involved in the turnover of the ECM.

A major species of protein secreted by these cells and involved in the turnover of the ECM are the metalloproteinases. Woessner, *FASEB J.* *5*:214-2154 (1991). As will be discussed below, new types of secretory glycoproteins have also been identified in human cartilage, osteoblasts, synovial cells, sheep and bovine oviduct and mammary cells, and macrophages. Nyrikos and Golds, *Biochem. J.* *268*:265-268 (1990); Hakala *et al.*, *J. Biol. Chem.268*(34):25803-25810 (1993); Johansen *et al.*, *J. Bone and Min. Res.* *7*(5):501-511 (1992); Rejman and Hurley, *Biochem. Biophys. Res. Commun.* *150*:329-334 (1988); DeSouza and Murray, *Endocrinology* *136*(6):2485-2496 (1995); Hollak *et al.*, *J. Clin. Invest.* *93*:1288-1292 (1995); Arias *et al.*, *Biol. of Reproduction* *51*:685-694 (1994). These proteins all share regions of significant homology to bacterial and fungal chitinases, and therefore are referred to herein as "chitinase-like proteins." They bear a subtle similarity to lysozymes from mammals and function as endoglycosidases, with a specificity for N-acetyl-glucosamine linkages. However, these types of chitin-like structures (homopolymers of N-acetyl-glucosamine) were not normally encountered in mammalian tissue.

The human cartilage glycoprotein, HC gp-39, is a mammalian chitinase-like protein with an apparent molecular weight of approximately 39 kDa that is secreted by both articular chondrocytes and synovial fibroblasts. Nyrikos and Golds, *supra*; Hakala *et al.*, *supra*. This protein has been described as a marker for joint injury, appearing in the blood and synovial fluid from patients diagnosed with rheumatoid arthritis. Johansen *et al.*, *British J. of Rheumatology 32*:949-955 (1993). The gene encoding this protein has been cloned. Hakala *et al.*, supra. Neither the protein nor mRNA for HC gp-39 was detectable in normal newborn or adult human articular cartilage obtained at surgery, but was detectable both in synovial specimens and in cartilage obtained from patients with rheumatoid arthritis. Hakala *et al.*, *supra*. The observations of Hakala *et al.* suggest that the expression of HC-gp-39 may be related to a response of these cells to an altered tissue environment.

The protein YKL-40 has also been identified as one of the major secretory products of cultured human osteoblastic cells (osteocarcinoma cell line MG-63) expressed in response to 1,25-dihydroxyvitamin D3 stimulation. Johansen *et al.*, *J. Bone and Min. Res. 7*(5):501-511 (1992); Johansen *et al.*, *British J. of Rheumatology 32*:949-955 (1993). The N-terminal portion of YKL-40 was sequenced and found to be identical to HC gp-39. Upon further sequencing, YKL-40 and HC gp-39 were found to be identical.

Chitotriosidase is an enzyme that has been identified as a member of the chitinase protein family. Renkema *et al.*, *J. Biol. Chem. 270*:2198-2202 (1995); Hollak *et al.*, *J. Clin. Invest. 93*:1288-1292 (1994). This protein also has an apparent molecular weight of 39 kDa and shares N-terminal homologies with HC gp-39, the bovine mammary protein, and several bacterial chitinases. Activity of this enzyme was originally detected from cells of patients afflicted with Gaucher Disease (GD). Gaucher Disease is an inherited deficiency in the activity of glucocerebrosidase, lysosomal hydrolase. The defect results in an accumulation of glucosylceramide (glucocerebroside) in the lysosomes of macrophages. Accumulation of lipid-laden macrophages results in hepatosplenomegaly, bone lesions, and neurological anomalies. After morphological differentiation of

monocytes in the macrophages in culture, the cells begin to produce secrete increasing amounts of chitotriosidase. This increase was, on average, 600 times greater in symptomatic GD patients than in healthy volunteers. Hollak *et al.*, supra. The elevation in chitotriosidase activity can be effectively reduced, however, upon initiation of enzyme supplementation therapy.

Unlike other members of the family of proteins homologous to chitinase, chitotriosidase does have chitinolytic activity. Renkema *et al.*, supra. Like the bacterial enzyme, it has ability to degrade chitin azure, a polymer of beta1-4-linked N-acetylglucosamine moieties. Chitotriosidase, along with HC gp-39, is believed to be involved in tissue remodeling in the mammalian cell, and thus may serve as useful tools in the development of therapeutics and diagnostics for tissue remodeling disorders. For example, the enzyme can serve as a valuable diagnostic marker for Gaucher disease, and may be useful in monitoring the efficacy of therapy.

Chitin is a major structural component of the cell walls of fungal pathogens. Chitinases are enzymes found in a variety of species, including plants, bacteria, fungi, insects, and fish (Flach *et al.*, *Experientia 48*:701-716 (1992)), that hydrolyze chitin or chitin-like substrates, such as 4-methylumbelliferyl chitotrioside. It is well documented that plant chitinases are potent inhibitors of fungal growth. Broekaert *et al.*, *Physiol. Mol. Plant Pathol. 33*:319-331 (1988); Roberts *et al.*, *J. Gen. Microbiol. 134*:169-174 (1988). Thus, plant chitinases may protect plants against fungi and other pathogens. It has often been suggested that humans and other mammals do not possess an analogous chitinase. Raghavan *et al.*, *Infect. Immun. 62*:1901-1908 (1994); Huber *et al.*, *Proc. Natl. Acad. Sci. USA 88*:2807-2810 (1991). This may not be the case.

Hollak *et al.*, discussed supra, has now observed that chitotriosidase is secreted in large quantities by activated human macrophages and has been reported to be markedly elevated in the plasma of patients suffering from Gaucher disease. Boot *et al.*, *J. Biol. Chem. 270*:26252-26256 (1995), has recently reported on the cloning of chitotriosidase. This report also suggests that this enzyme may play a role in the degradation of chitin-containing pathogens. Escott

and Adams, *Infect. Immun.* 63:4770-4773 (1995), report on the detection of a chitinase in human serum and leukocytes that cleaves colloidal [^{3}H] chitin. They report that the chitinase activity detected in human granulocytes may reflect that fact that the enzyme in humans may play a protective role against fungi and other chitin-containing microbial pathogens. This role is conceivable in view of the well-documented role of plant chitinases as potent inhibitors of fungal growth.

Accordingly, there is a need in the art for the identification of novel human chitinase enzymes, which along with chitotriosidase and HC gp-39, can serve as useful tools in the development of therapeutics and diagnostics for tissue remodeling disorders or as anti-fungal agents.

Summary of the Invention

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding a human chitinase alpha polypeptide having the amino acid sequence in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4] or the amino acid sequence encoded by the cDNA clone deposited in a bacterial host as ATCC Deposit Number 97529 on May 3, 1996. The present invention further provides isolated nucleic acid molecules comprising a polynucleotide encoding a human chitinase alpha-2 polypeptide having the amino acid sequence in Figure 5 [SEQ ID NO:12].

The nucleotide sequence determined by sequencing the deposited chitinase alpha cDNA clone, which is shown in Figure 1 [SEQ ID NO:1], contains an open reading frame encoding a polypeptide of about 421 amino acid residues including an initiation codon at nucleotide positions 82-84, a leader sequence of about 26 amino acid residues and a deduced molecular weight of about 44 kDa. The 395 amino acids of the predicted mature chitinase alpha protein is shown in Figure 1 (last 395 residues) and in SEQ ID NO:2 (from amino acid residue 27 to residue 421).

The nucleotide sequence shown in Figure 5 [SEQ ID NO:11] contains an open reading frame encoding a polypeptide of about 385 amino acid residues

including an initiation codon at nucleotide positions 27-29 and a leader sequence of about 20 amino acid residues. The 365 amino acids of the predicted mature chitinase alpha-2 protein is shown in Figure 5 (last 365 residues) and in SEQ ID NO:12 (from amino acid residue 21 to residue 385). As discussed in more detail below, the chitinase alpha-2 gene is a shorter variant of the chitinase alpha gene.

Thus, one aspect of the invention provides isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding the chitinase alpha polypeptide having the complete amino acid sequence shown in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4]; (b) a nucleotide sequence encoding the mature chitinase alpha polypeptide having the amino acid sequence at positions 27-421 in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4]; (c) a nucleotide sequence encoding the chitinase alpha polypeptide having the complete amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97529; (d) a nucleotide sequence encoding the mature chitinase alpha polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97529; (e) a nucleotide sequence encoding the chitinase alpha-2 polypeptide having the complete amino acid sequence shown in Figure 5 [SEQ ID NO:12]; (f) a nucleotide sequence encoding the mature chitinase alpha-2 polypeptide having the amino acid sequence at positions 21-385 in Figure 5 [SEQ ID NO:12]; and (g) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e) or (f) above. Preferably, the nucleic acid molecule will encode the mature polypeptide as shown in Figure 1 [SEQ ID NO:2], Figure 2 [SEQ ID NO:4], Figure 5 [SEQ ID NO:12] or as encoded by the above-described deposited cDNA.

Further embodiments of the invention include isolated nucleic acid molecules that comprise a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 97%, 98% or 99% identical, to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), or (g) above, or a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide having a nucleotide sequence identical to a nucleotide sequence in

(a), (b), (c), (d), (e), (f) or (g) above. The polynucleotide which hybrizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues. An additonal nucleic acid embodiment of the invention relates to an isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of an chitinase alpha polypeptide having an amino acid sequence in (a), (b), (c), (d), (e), or (f) above.

The present invention also relates to recombinant vectors which include the isolated nucleic acid molecules of the present invention and host cells containing the recombinant vectors, as well as methods of making such vectors and host cells and for using them for production of chitinase alpha- and chitinase alpha-2- polypeptides or peptides by recombinant techniques.

The invention further provides an isolated chitinase alpha or chitinase alpha-2 polypeptide having amino acid sequence selected from the group consisting of: (a) the amino acid sequence of the chitinase alpha polypeptide having the complete 421 amino acid sequence including the leader sequence shown in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4]; (b) the amino acid sequence of the mature chitinase alpha polypeptide (without the leader) having the amino acid sequence at positions 27-421 in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4]; (c) the amino acid sequence of the chitinase alpha polypeptide having the complete amino acid sequence including the leader encoded by the cDNA clone contained in ATCC Deposit No. 97529; (d) the amino acid sequence of the mature chitinase alpha polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97529; (e) the amino acid sequence of the chitinase alpha-2 polypeptide having the complete 385 amino acid sequence including the leader sequence shown in Figure 5 [SEQ ID NO:12]; and (f) the amino acid sequence of the mature chitinase alpha-2 polypeptide (without the leader) having the amino acid sequence at positions 21-385 in Figure 5 [SEQ ID NO:12].

The polypeptides of the present invention also include polypeptides having an amino acid sequence with at least 90% similarity, more preferably at least 95%

similarity to those described in (a), (b), (c), (d), (e), or (f) above, as well as polypeptides having an amino acid sequence at least 80% identical, more preferably at least 90% identical, and still more preferably 95%, 97%, 98% or 99% identical to those above.

An additional embodiment of this aspect of the invention relates to a peptide or polypeptide which has the amino acid sequence of an epitope-bearing portion of a chitinase alpha polypeptide or a chitinase alpha-2 polypeptide having an amino acid sequence described in (a), (b), (c), (d), (e), or (f) above. Peptides or polypeptides having the amino acid sequence of an epitope-bearing portion of a chitinase alpha polypeptide or a chitinase alpha-2 polypeptide of the invention include portions of such polypeptides with at least six or seven, preferably at least nine, and more preferably at least about 30 amino acids to about 50 amino acids, although epitope-bearing polypeptides of any length up to and including the entire amino acid sequence of a polypeptide of the invention described above also are included in the invention. In another embodiment the invention provides an isolated antibody that binds specifically to a chitinase alpha polypeptide or chitinase alpha-2 polypeptide having an amino acid sequence described in (a), (b), (c), (d), (e), or (f) above.

Chitinase alpha and chitinase alpha-2, which are secreted, and which have significant homology to human cartilage glycoprotein-39 (HC gp-39) and to chitotriosidase, are believed by the present inventors to be involved in tissue remodeling. Therefore, chitinase alpha and chitinase alpha-2 are useful targets for the inhibition of tissue remodeling disease processes, such as rheumatoid arthritis. Accordingly, the invention is related to a method of treating an individual in need of a decreased level of chitinase alpha or chitinase alpha-2 activity in the body, which involves administering to such an individual a composition comprising an antagonist to chitinase alpha and/or chitinase alpha-2 such as anti-chitinase alpha- or anti-chitinase alpha-2-antibodies.

It is also believed by the inventors that the chitinase alpha polypeptide and the chitinase alpha-2 polypeptide of the invention can be used to treat fungal infection, as chitin is a major component of the fungal cell wall, and chitnases

cleave this polymer. Thus, an additional aspect of the invention is related to a method for treating fungal infection, which involves administering to such an indivdual a composition comprising a chitinase alpha polypeptide or a chitinase alpha-2 polypeptide of the invention.

Further, for a number of disorders, it is believed by the inventors that significantly higher or lower levels of chitinase alpha and/or chitinase alpha-2 gene expression can be detected in bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to a "standard" chitinase alpha or chitinase alpha-2 gene expression level, i.e., the chitinase alpha or chitinase alpha-2 gene expression level in bodily fluids from an individual not having the disorder. Thus, the invention provides a diagnostic method useful during diagnosis of a disorder related to an abnormal level of chitinase alpha or chitinase alpha-2 gene expression, which involves (a) assaying chitinase alpha or chitinase alpha-2 gene expression level in cells or body fluid of that individual; (b) comparing that chitinase alpha or chitinase alpha-2 gene expression level with a standard chitinase alpha or chitinase alpha-2 gene expression level, whereby an increase or decrease in the assayed chitinase alpha or chitinase alpha-2 gene expression level compared to the standard expression level is indicative of a disorder.

Brief Description of the Figures

Figure 1 provides the coding sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) of the polypeptide chitinase alpha.

Figure 2 provides the coding sequence (SEQ ID NO:3) and deduced amino acid sequence (SEQ ID NO:4) of a chitinase alpha variant. The nucleotide sequence is modified from Figure 1 at positions 607-609.

Figure 3 shows regions of similarity between the amino acid sequences of chitinase alpha (HPMFW51) (SEQ ID NO:4), HC gp-39 (SEQ ID NO:5), and chitotriosidase (SEQ ID NO:6).

Figure 4 shows an analysis of the chitinase alpha amino acid sequence. In particular, the Kyte-Doolittle hydrophilicity plot and the Jameson-Wolf antigenic inex are provided. Amino acid residues Leu38-Ala51; Asn66-Phe85; Asn117-Lys127; Met145-Leu153; Asn165-Asp172; Tyr177-His185; Phe245-Ile254; Gly256-Tyr274; Lys286-Tyr293; Gly366-Lys376; and Asp394-Pro407 in Figure 1 are highly antigenic regions of the chitinase alpha protein.

Figure 5 provides the coding sequence (SEQ ID NO:11) and deduced amino acid sequence (SEQ ID NO:12) of a shortened chitinase alpha variant – chitinase alpha-2.

Figures 6(a-d) provide a comparison of the nucleotide (a-c) and deduced amino acid (d) sequences of chitinase alpha-2 and chitinase alpha.

Detailed Description of the Invention

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding the chitinase alpha protein having the amino acid sequence shown in Figure 1 [SEQ ID NO:2] which was determined by sequencing a cloned cDNA. The nucleotide sequence shown in Figure 1 [SEQ ID NO:1] was obtained by sequencing the hpmfw51 cDNA clone encoding a chitinase alpha polypeptide, which was deposited on May 3 , 1996 at the American Type Culture Collection, 12301 Park Lawn Drive, Rockville, Maryland 20852, and given accession number 97529. The deposited clone is contained in the pBluescript SK(-) plasmid (Stratagene, LaJolla, CA).

The present invention further provides isolated nucleic acid molecules comprising a polynucleotide encoding the chitinase alpha-2 protein having the amino acid sequence shown in Figure 5 [SEQ ID NO:12].

Nucleic Acid Molecules

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer (such as the Model 373 from Applied Biosystems, Inc.), and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain a some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

Unless otherwise indicated, each "nucleotide sequence" set forth herein is presented as a sequence of deoxyribonucleotides (abbreviated A, G , C and T). However, by "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U) where each thymidine deoxynucleotide (T) in the specified deoxynucleotide sequence in is replaced by the ribonucleotide uridine (U). For instance, reference to an RNA molecule having the sequence of SEQ ID NO:1 set forth using deoxyribonucleotide abbreviations is intended to indicate an RNA molecule having a sequence in which each deoxynucleotide A, G or C of SEQ ID NO:1 has been replaced by the

corresponding ribonucleotide A, G or C, and each deoxynucleotide T has been replaced by a ribonucleotide U.

Using the information provided herein, such as the nucleotide sequence in Figure 1 [SEQ ID NO:1], Figure 2 [SEQ ID NO:3], or Figure 5 [SEQ ID NO:11], a nucleic acid molecule of the present invention encoding a chitinase alpha or chitinase alpha-2 polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Illustrative of the invention, the nucleic acid molecule described in Figure 1 [SEQ ID NO:1] was discovered in a cDNA library derived from human placenta tissue. The determined nucleotide sequence of the chitinase alpha cDNA of Figure 1 contains an open reading frame encoding a protein of about 421 amino acid residues with an initiation codon at positions 82-84 of the nucleotide sequence shown in Figure 1 [SEQ ID NO. 1], and a predicted leader sequence of about 26 amino acid residues, and a deduced molecular weight of about 44 kDa without the leader and about 46.6 kDa with the leader. The amino acid sequence of the predicted mature chitinase alpha protein is also shown in Figure 1 [SEQ ID NO:2] from about amino acid residue 27 to about residue 421. The chitinase alpha protein shown in Figure 1 [SEQ ID NO:2] is about 52% identical and about 68% similar to human gp-39, and about 52% identical and about 70% similar to chitotriosidase (Figure 3).

The determined nucleotide sequence of the chitinase alpha-2 cDNA of Figure 5 contains an open reading frame encoding a protein of about 385 amino acid residues with an initiation codon at positions 27-29 of the nucleotide sequence shown in Figure 5 [SEQ ID NO: 11], and a predicted leader sequence of about 20 amino acid residues. The amino acid sequence of the predicted mature chitinase alpha-2 protein is also shown in Figure 5 [SEQ ID NO:12] from about amino acid residue 21 to about residue 385. The chitinase alpha-2 nucleotide sequence shown in Figure 5 [SEQ ID NO:11] is about 99% identical and about 99 % similar to the chitinase alpha nucleotide sequence shown in Figure 1 [SEQ ID NO:1].

As one of ordinary skill would appreciate, due to the possibilities of sequencing errors discussed above, as well as the variability of cleavage sites for leaders in different known proteins, the actual chitinase alpha polypeptide encoded by the deposited cDNA comprises about 421 amino acids, but may be anywhere in the range of 410-432 amino acids; and the actual leader sequence of this protein is about 26 amino acids, but may be anywhere in the range of about 21 to about 31 amino acids. Similarly, the actual chitinase alpha-2 polypeptide comprises about 385 amino acids, but may be anywhere in the range of 375-395 amino acids; and the actual leader sequence of this protein is about 20 amino acids, but may be anywhere in the range of about 16 to about 25 amino acids.

As indicated, nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

Isolated nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) with an initiation codon at positions 82-84 of the nucleotide sequence shown in Figure 1 [SEQ ID NO:1] or Figure 2 [SEQ ID NO:3]; DNA molecules comprising the coding sequence for the mature chitinase alpha protein shown in Figure 1 (last 395 amino acids) [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4] (last 395 amino acids); DNA molecules

comprising an open reading frame (ORF) with an initiation codon at positions 27-29 of the nucleotide sequence shown in Figure 5 [SEQ ID NO:11]; DNA molecules comprising the coding sequence for the mature chitinase alpha-2 protein shown in Figure 5 (last 365 amino acids) [SEQ ID NO:12]; and DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode the chitinase alpha protein. Of course, the genetic code is well known in the art. Thus, it would be routine for one skilled in the art to generate the degenerate variants described above.

In another aspect, the invention provides isolated nucleic acid molecules encoding the chitinase alpha polypeptide having an amino acid sequence encoded by the cDNA clone contained in the plasmid deposited as ATCC Deposit No 97529 on May 3, 1996. Preferably, this nucleic acid molecule will encode the mature polypeptide encoded by the above-described deposited cDNA clone. The invention further provides an isolated nucleic acid molecule having the nucleotide sequence shown in Figure 1 [SEQ ID NO:1] or Figure 2 [SEQ ID NO:3] or Figure 5 [SEQ ID NO:11] or the nucleotide sequence of the chitinase alpha cDNA contained in the above-described deposited clone, or nucleic acid molecule having a sequence complementary to one of the above sequences. Such isolated molecules, particularly DNA molecules, are useful as probes for gene mapping by *in situ* hybridization with chromosomes and for detecting expression of the chitinase alpha gene in human tissue, for instance, by Northern blot analysis.

The present invention is further directed to fragments of the isolated nucleic acid molecules described herein. By a fragment of an isolated nucleic acid molecule having the nucleotide sequence of the deposited cDNA or the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 2 (SEQ ID NO:3), or Figure 5 (SEQ ID NO:11), is intended fragments at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length which are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments 50-1400 nt in length are also useful according to the present invention as are fragments

corresponding to most, if not all, of the nucleotide sequence of the deposited cDNA or as shown in Figure 1 (SEQ ID NO:1), Figure 2 (SEQ ID NO:3), or Figure 5 (SEQ ID NO:11). By a fragment at least 20 nt in length, for example, is intended fragments which include 20 or more contiguous bases from the nucleotide sequence of the deposited cDNA or the nucleotide sequence as shown in Figure 1 (SEQ ID NO:1), Figure 2 (SEQ ID NO:3), or Figure 5 (SEQ ID NO:11). Since the gene has been deposited and the nucleotide sequences shown in Figure 1 (SEQ ID NO:1), Figure 2 (SEQ ID NO:3) and Figure 5 (SEQ ID NO:11) are provided, generating such DNA fragments would be routine to the skilled artisan. For example, restriction endonuclease cleavage or shearing by sonication could easily be used to generate fragments of various sizes. Alternatively, such fragments could be generated synthetically.

Preferred nucleic acid fragments of the present invention include nucleic acid molecules encoding epitope-bearing portions of the chitinase alpha and chitinase alpha-2 proteins. In particular, such nucleic acid fragments of the present invention include nucleic acid molecules encoding: a polypeptide comprising amino acid residues from about 38 to about 51 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 66 to about 85 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 117 to about 127 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 145 to about 153 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 165 to about 172 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 177 to about 185 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 245 to about 254 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 256 to about 274 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 286 to about 293 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 366 to about 376 in Figure 1 (SEQ ID NO:2); and a polypeptide comprising amino acid residues from about 394 to about 407 in Figure 1 (SEQ ID NO:2).

The inventors have determined that the above polypeptide fragments are antigenic regions of the chitinase alpha protein. Methods for determining other such epitope-bearing portions of the chitinase alpha protein are described in detail below.

In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention described above, for instance, the cDNA clone contained in ATCC Deposit 97529. By "stringent hybridization conditions" is intended overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150 mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C. By a polynucleotide which hybridizes to a "portion" of a polynucleotide is intended a polynucleotide (either DNA or RNA) hybridizing to at least about 15 nucleotides (nt), and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably about 30-70 nt of the reference polynucleotide. These are useful as diagnostic probes and primers as discussed above and in more detail below.

Of course, polynucleotides hybridizing to a larger portion of the reference polynucleotide (e.g., the deposited cDNA clone), for instance, a portion 50-750 nt in length, or even to the entire length of the reference polynucleotide, also useful as probes according to the present invention, as are polynucleotides corresponding to most, if not all, of the nucleotide sequence of the deposited cDNA or the nucleotide sequence as shown in Figure 1 [SEQ ID NO:1], Figure 2 [SEQ ID NO:3] or Figure 5 [SEQ ID NO:11]. By a portion of a polynucleotide of "at least 20 nt in length," for example, is intended 20 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide, (e.g., the deposited cDNA or the nucleotide sequence as shown in Figure 1 [SEQ ID NO:1], Figure 2.[SEQ ID NO:3] or Figure 5 [SEQ ID NO:11]). As indicated, such portions are useful diagnostically either as a probe according to conventional

DNA hybridization techniques or as primers for amplification of a target sequence by the polymerase chain reaction (PCR), as desribed, for instance, in *Molecular Cloning, A Laboratory Manual*, 2nd. edition, Sambrook, J. *et al.*, eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989), the entire disclosure of which is hereby incorporated herein by reference.

Since a chitinase alpha cDNA clone has been deposited and its determined nucleotide sequence is provided in Figure 1 [SEQ ID NO:1] and the nucleotide sequence of a variant is provided in Figure 2 [SEQ ID NO:3], generating polynucleotides which hybridize to a portion of the chitinase alpha cDNA molecule would be routine to the skilled artisan. Similarly, since the determined nucleotide sequence for the chitinase alpha-2 protein is provided in Figure 5 [SEQ ID NO:11], generating polynucleotides which hybridize to a portion of the chitinase alpha-2 cDNA molecule would be routine to the skilled artisan. For example, restriction endonuclease cleavage or shearing by sonication of the chitinase alpha cDNA clone or chitinase alpha-2 cDNA clone could easily be used to generate DNA portions of various sizes which are polynucleotides that hybridize to a portion of the chitinase alpha or chitinase alpha-2 cDNA molecule. Alternatively, the hybridizing polynucleotides of the present invention could be generated synthetically according to known techniques. Of course, a polynucleotide which hybridizes only to a poly A sequence (such as the 3′ terminal poly(A) tract of the chitinase alpha cDNA shown in Figure 1 [SEQ ID NO:1]) or Figure 2 [SEQ ID NO:3], or to a complementary stretch of T (or U) resides, would not be included in a polynucleotide of the invention used to hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule contain a poly (A) stretch or the complement thereof (e.g., practially any double-standed cDNA clone).

As indicated, nucleic acid molecules of the present invention which encode the chitinase alpha or chitinase alpha-2 polypeptide may include, but are not limited to those encoding the amino acid sequence of the mature polypeptide, by itself; the coding sequence for the mature polypeptide and additional sequences, such as those encoding the about 26 amino acid leader or secretory sequence, such

as a pre-, or pro- or prepro- protein sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing - including splicing and polyadenylation signals, for example - ribosome binding and stability of mRNA; an additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities. Thus, the sequence encoding the polypeptide may be fused to a marker sequence, such as a sequence encoding a peptide which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (Qiagen, Inc.), among others, many of which are commercially available. As described in Gentz *et al.*, *Proc. Natl. Acad. Sci. (USA) 86*:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. The "HA" tag is another peptide useful for purification which corresponds to an epitope derived from the influenza hemagglutinin protein, which has been described by Wilson *et al.*, *Cell 37*:767 (1984).

The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the chitinase alpha or chitinase alpha-2 protein. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an orgnanism. *Genes II*, Lewin, ed. Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

Such variants include those produced by nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties

and activities of the chitinase alpha or chitinase alpha-2 protein or portions thereof. Also especially preferred in this regard are conservative substitutions.

Further embodiments of the invention include isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical to (a) a nucleotide sequence encoding the full-length chitinase alpha polypeptide having the complete amino acid sequence in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4] including the predicted leader sequence; (b) a nucleotide sequence encoding the mature chitinase alpha polypeptide (full-length polypeptide with the leader removed) having the amino acid sequence at positions 27-421 in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4]; a nucleotide sequence encoding the full-length chitinase alpha polypeptide having the complete amino acid sequence including the leader encoded by the cDNA clone contained in ATCC Deposit No. 97529; (d) a nucleotide sequence encoding the mature chitinase alpha polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97529; (e) a nucleotide sequence encoding the chitinase alpha-2 polypeptide having the complete amino acid sequence shown in Figure 5 [SEQ ID NO:12]; (f) a nucleotide sequence encoding the mature chitinase alpha-2 polypeptide having the amino acid sequence at positions 21-385 in Figure 5 [SEQ ID NO:12]; and (g) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e), or (f).

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding an chitinase alpha or chitinase alpha-2 polypeptide is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the chitinase alpha or chitinase alpha-2 polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides

in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5′ or 3′ terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular nucleic acid molecule is at least 90%, 95%, 97%, 98% or 99% identical to, for instance, the nucleotide sequence shown in Figure 1, Figure 2, or Figure5 or to the nucleotides sequence of the deposited cDNA clone can be determined conventionally using known computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). Bestfit uses the local homology algorithm of Smith and Waterman (*Advances in Applied Mathematics 2*:482-489 (1981)) to find the best segment of homology between two sequences. When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed.

The present application is directed to nucleic acid molecules at least 90%, 95%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 [SEQ ID NO:1] or Figure 2 [SEQ ID NO:3] or Figure 5 [SEQ ID NO:11] or to the nucleic acid sequence of the deposited cDNA, irrespective of whether they encode a polypeptide having chitinase alpha or chitinase alpha-2 activity. This is because, even where a particular nucleic acid molecule does not encode a polypeptide having such activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having chitinase alpha or chitinase alpha-2 activity include, *inter alia*, (1) isolating the chitinase alpha or

chitinase alpha-2 gene or allelic variants thereof in a cDNA library; (2) *in situ* hybridization (e.g., "FISH") to metaphase chromosomal spreads to provide precise chromosomal location of the chitinase alpha or chitinase alpha-2 gene as described in Verma *et al.*, *Human Chromosomes: a Manual of Basic Techniques*, Pergamon Press, New York (1988); and Northern Blot analysis for detecting chitinase alpha or chitinase alpha-2 mRNA expression in specific cell types (e.g., placenta).

Preferred, however, are nucleic acid molecules having sequences at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 [SEQ ID NO:1] or to the nucleic acid sequence of the deposited cDNA which do, in fact, encode a polypeptide having chitinase alpha or chitinase alpha-2 protein activity. By "a polypeptide having chitinase alpha or chitinase alpha-2 activity" is intended polypeptides exhibiting activity similar, but not necessarily identical, to an activity of the chitinase alpha or chitinase alpha-2 protein of the invention (either the full-length protein or, preferably, the mature protein) as measured in a particular biological assay. One such assay for chitinase activity is the assay described by Dickinson *et al.*, *J. Gen. Microbiol. 135*:1417-1421 (1989).

Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 90%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequence of the deposited cDNA or the nucleic acid sequence shown in Figure 1 [SEQ ID NO:1] or Figure 2 [SEQ ID NO:3] or Figure 5 [SEQ ID NO:11] will encode a polypeptide "having chitinase alpha or chitinase alpha-2 protein activity." In fact, since degenerate variants of these nucleotide sequences all encode the same polypeptide, this will be clear to the skilled artisan even without performing one of the above-described comparison assays. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having chitinase alpha or chitinase alpha-2 protein activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either

less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid).

For example, guidance concerning how to make phenotypically silent amino acid substitions is provided in Bowie, J.U. *et al.*, "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," *Science 247:*1306-1310 (1990), wherein the authors indicate that there are two main approaches for studying the tolerance of an amino acid sequence to change. The first method relies on the process of evolution, in which mutations are either accepted or rejected by natural selection. The second approach uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene and selections or screens to identify sequences that maintain functionality. As the authors state, these studies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at a certain position of the protein. For example, most buried amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Other such phenotypically silent substitutions are described in Bowie, J.U., *et al., supra,* and the references cited therein.

Thus, in accordance with another aspect of the invention, the following nucleotide sequence variations of the chitinase alpha gene in Figure 1 [SEQ ID NO:1] are provided: the nucleotides at positions 607-609 in Figure 1 [SEQ ID NO:1] are changed to either GAA or GAG. This results in an amino acid substitution of glutamic acid for isoleucine. Figure 2 [SEQ ID NO:3] depicts this sequence variant, which is expected to encode a peptide having chitinolytic activity.

Vectors and Host Cells

The present invention also relates to vectors which include the isolated DNA molecules of the present invention, host cells which are genetically engineered with the recombinant vectors, and the production of chitinase alpha polypeptides or portions thereof by recombinant techniques.

Recombinant constructs may be introduced into host cells using well known techniques such as infection, transduction, transfection, transvection, electroporation and transformation. The vector may be, for example, a phage, plasmid, viral or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged *in vitro* using an appropriate packaging cell line and then transduced into host cells.

Preferred are vectors comprising cis-acting control regions to the polynucleotide of interest. Appropriate trans-acting factors may be supplied by the host, supplied by a complementing vector or supplied by the vector itself upon introduction into the host.

In certain preferred embodiments in this regard, the vectors provide for specific expression, which may be inducible and/or cell type-specific. Particularly preferred among such vectors are those inducible by environmental factors that are easy to manipulate, such as temperature and nutrient additives.

Expression vectors useful in the present invention include chromosomal-, episomal- and virus-derived vectors, e.g., vectors derived from bacterial plasmids, bacteriophage, yeast episomes, yeast chromosomal elements, viruses such as baculoviruses, papova viruses, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as cosmids and phagemids.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac*, *trp* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation,termination and, in the transcribed region, a ribosome binding site for translation. The coding portion

of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture and tetracycline or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include bacterial cells, such as *E. coli*, Streptomyces and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Appropriate culture media and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; pSVK3, pBPV, pMSG and pSVL available from Pharmacia, and pA2 available from Qiagen. Other suitable vectors will be readily apparent to the skilled artisan.

Among known bacterial promoters suitable for use in the present invention include the *E. coli lac*I and *lac*Z promoters, the T3 and T7 promoters, the *gpt* promoter, the lambda PR and PL promoters and the *trp* promoter. Suitable eukaryotic promoters include the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus (RSV), and metallothionein promoters, such as the mouse metallothionein-I promoter.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods.

Such methods are described in many standard laboratory manuals, such as Davis *et al.*, *Basic Methods in Molecular Biology* (1986).

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. The signals may be endogenous to the polypeptide or they may be heterologous signals.

Thus, the polypeptides of the invention may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptides to improve stability and persistence in the host cell, during purification or during subsequent handling and storage. Also, region(s) also may be added to the polypeptides to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art. A preferred fusion protein comprises a heterologous region from immunoglobulin that is useful to solubilize polypeptides. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobin molecules together with another human protein or part thereof. In many cases, the Fc part in fusion protein is thoroughly advantageous for use in therapy and diagnosis and thus

results, for example, in improved pharmacokinetic properties (EP-A 0232 262). On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified in the advantageous manner described. This is the case when Fc portion proves to be a hindrance to use in therapy and diagnosis, for example when the fusion protein is to be used as antigen for immunizations. In drug discovery, for example, human proteins, such as, hIL5- has been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. See, Bennett, D. *et al.*, *Journal of Molecular Recognition* 8:52-58 (1995) and Johanson, K. *et al.*, *The Journal of Biological Chemistry* 270(16):9459-9471 (1995).

The chitinase alpha or chitinase alpha-2 protein can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

Chitinase Alpha and Chitinase Alpha-2 Polypeptides and Peptides

The invention further provides an isolated chitinase alpha or chitinase alpha-2 polypeptide having the amino acid sequence encoded by the deposited cDNA, or the amino acid sequence in Figure 1 [SEQ ID NO:2], Figure 2 [SEQ

ID NO:4], or Figure 5 [SEQ ID NO:12] or a peptide or polypeptide comprising a portion of the above polypeptides. The terms "peptide" and "oligopeptide" are considered synonymous (as is commonly recognized) and each term can be used interchangeably as the context requires to indicate a chain of at least to amino acids coupled by peptidyl linkages. The word "polypeptide" is used herein for chains containing more than ten amino acid residues. All oligopeptide and polypeptide formulas or sequences herein are written from left to right and in the direction from amino terminus to carboxy terminus.

It will be recognized in the art that some amino acid sequence of the chitinase alpha or chitinase alpha-2 polypeptide can be varied without significant effect of the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine activity. In general, it is possible to replace residues which form the tertiary structure, provided that residues performing a similar function are used. In other instances, the type of residue may be completely unimportant if the alteration occurs at a non-critical region of the protein.

Thus, the invention further includes variations of the chitinase alpha and chitinase alpha-2 polypeptides which show substantial chitinase alpha or chitinase alpha-2 polypeptide activity. Such mutants include deletions, insertions, inversions, repeats, and type substitutions (for example, substituting one hydrophilic residue for another, but not strongly hydrophilic for strongly hydrophobic as a rule). Small changes or such "neutral" amino acid substitutions will generally have little effect on activity.

Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

A particularly preferred variant of the chitinase alpha in Figure 1 [SEQ ID NO:2] are those which result in an amino acid substitution of glutamic acid for isoleucine at amino acid position 176 (i.e., the nucleotides at positions 607 to 609

in Figure 1 are changed to either GAA or GAG). Figure 2 [SEQ ID NO:4] depicts this preferred sequence variant, which is expected to encode a peptide having chitinolytic activity .

As indicated in detail above, further guidance concerning which amino acid changes are likely to be phenotypically silent (i.e., are not likely to have a significant deleterious effect on a function) can be found in Bowie, J.U., *et al.*, *Science 247:*1306-1310 (1990).

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of the chitinase alpha or chitinase alpha-2 polypeptide can be substantially purified by the one-step method described in Smith and Johnson, *Gene 67:*31-40 (1988).

The polypeptides of the present invention include the polypeptide encoded by the deposited cDNA including the leader, the mature polypeptide encoded by the deposited the cDNA minus the leader (i.e., the mature protein), the polypeptide of Figure 1 [SEQ ID NO:2], Figure 2 [SEQ ID NO:4] or Figure 5 [SEQ ID NO:12] including the leader, the polypeptide of Figure 1 [SEQ ID NO:2], Figure 2 [SEQ ID NO:4] or Figure 5 [SEQ ID NO:12] minus the leader, as well as polypeptides which have at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 97%, 98% or 99% similarity to those described above. Further polypeptides of the present invention include polypeptides at least 80% identical, more preferably at least 90% or 95% identical, still more preferably at least 97%, 98% or 99% identical to the polypeptide encoded by the deposited cDNA, to the polypeptide of SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:12 and also include portions of such polypeptides with at least 30 amino acids and more preferably at least 50 amino acids.

By "% similarity" for two polypeptides is intended a similarity score produced by comparing the amino acid sequences of the two polypeptides using the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711) and the default settings for determining similarity. Bestfit

uses the local homology algorithm of Smith and Waterman (*Advances in Applied Mathematics 2*:482-489 (1981)) to find the best segment of similarity between two sequences.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a reference amino acid sequence of a chitinase alpha or chitinase alpha-2 polypeptide is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of the chitinase alpha polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 97%, 98% or 99% identical to, for instance, the amino acid sequence shown in Figure 1 [SEQ ID NO:2], Figure 2 [SEQ ID NO:4] or Figure 5 [SEQ ID NO:12], or to the amino acid sequence encoded by deposited cDNA clone, can be determined conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed.

As described in detail below, the polypeptides of the present invention can be used to raise polyclonal and monoclonal antibodies, which are useful in diagnostic assays for detecting chitinase alpha and/or chitinase alpha-2 protein expression as described below or as agonists and antagonists capable of enhancing or inhibiting protein function. Further, such polypeptides can be used in the yeast two-hybrid system to "capture" chitinase alpha or chitinase alpha-2 protein binding proteins which are also candidate agonist and antagonist according to the present invention. The yeast two hybrid system is described in Fields and Song, *Nature 340:*245-246 (1989).

In another aspect, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. These immunogenic epitopes are believed to be confined to a few loci on the molecule. On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes. See, for instance, Geysen, H.M. *et al., Proc. Natl. Acad. Sci. (USA) 81:*3998-4002 (1984).

As to the selection of peptides or polypeptides bearing an antigenic epitope (i.e., that contain a region of a protein molecule to which an antibody can bind), it is well known in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein. See, for instance, Sutcliffe, J.G. *et al., Science 219:*660-666 (1983). Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins (i.e., immunogenic epitopes) nor to the amino or carboxyl terminals. Peptides that are extremely hydrophobic and those of six or fewer residues generally are ineffective at inducing antibodies that bind

to the mimicked protein; longer, soluble peptides, especially those containing proline residues, usually are effective. Sutcliffe *et al.*, *supra*, at 661. For instance, 18 of 20 peptides designed according to these guidelines, containing 8-39 residues covering 75% of the sequence of the influenza virus hemagglutinin HA1 polypeptide chain, induced antibodies that reacted with the HA1 protein or intact virus; and 12/12 peptides from the MuLV polymerase and 18/18 from the rabies glycoprotein induced antibodies that precipitated the respective proteins.

Antigenic epitope-bearing peptides and polypeptides of the invention are therefore useful to raise antibodies, including monoclonal antibodies, that bind specifically to a polypeptide of the invention. Thus, a high proportion of hybridomas obtained by fusion of spleen cells from donors immunized with an antigen epitope-bearing peptide generally secrete antibody reactive with the native protein. Sutcliffe *et al.*, *supra*, at 663. The antibodies raised by antigenic epitope-bearing peptides or polypeptides are useful to detect the mimicked protein, and antibodies to different peptides may be used for tracking the fate of various regions of a protein precursor which undergoes posttranslation processing. The peptides and anti-peptide antibodies may be used in a variety of qualitative or quantitative assays for the mimicked protein, for instance in competition assays since it has been shown that even short peptides (e.g., about 9 amino acids) can bind and displace the larger peptides in immunoprecipitation assays. See, for instance, Wilson, I.A. *et al.*, *Cell 37*:767-778 (1984) at 777. The anti-peptide antibodies of the invention also are useful for purification of the mimicked protein, for instance, by adsorption chromatography using methods well known in the art.

Antigenic epitope-bearing peptides and polypeptides of the invention designed according to the above guidelines preferably contain a sequence of at least seven, more preferably at least nine and most preferably between about 15 to about 30 amino acids contained within the amino acid sequence of a polypeptide of the invention. However, peptides or polypeptides comprising a larger portion of an amino acid sequence of a polypeptide of the invention, containing about 30 to about 50 amino acids, or any length up to and including the entire amino acid sequence of a polypeptide of the invention, also are considered

epitope-bearing peptides or polypeptides of the invention and also are useful for inducing antibodies that react with the mimicked protein. Preferably, the amino acid sequence of the epitope-bearing peptide is selected to provide substantial solubility in aqueous solvents (i.e., the sequence includes relatively hydrophilic residues and highly hydrophobic sequences are preferably avoided); and sequences containing proline residues are particularly preferred.

Non-limiting examples of antigenic polypeptides or peptides that can be used to generate chitinase alpha- or chitinase alpha-2-specific antibodies include: a polypeptide comprising amino acid residues from about 38 to about 51 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 66 to about 85 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 117 to about 127 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 145 to about 153 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 165 to about 172 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 177 to about 185 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 245 to about 254 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 256 to about 274 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 286 to about 293 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 366 to about 376 in Figure 1 (SEQ ID NO:2); and a polypeptide comprising amino acid residues from about 394 to about 407 in Figure 1 (SEQ ID NO:2).

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means for making peptides or polypeptides including recombinant means using nucleic acid molecules of the invention. For instance, a short epitope-bearing amino acid sequence may be fused to a larger polypeptide which acts as a carrier during recombinant production and purification, as well as during immunization to produce anti-peptide antibodies. Epitope-bearing peptides also may be synthesized using known methods of chemical synthesis. For instance, Houghten has described a simple method for

synthesis of large numbers of peptides, such as 10-20 mg of 248 different 13 residue peptides representing single amino acid variants of a segment of the HA1 polypeptide which were prepared and characterized (by ELISA-type binding studies) in less than four weeks. Houghten, R.A., *Proc. Natl. Acad. Sci. (USA)* *82*:5131-5135 (1985). This "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten *et al.* (1986). In this procedure the individual resins for the solid-phase synthesis of various peptides are contained in separate solvent-permeable packets, enabling the optimal use of the many identical repetitive steps involved in solid-phase methods. A completely manual procedure allows 500-1000 or more syntheses to be conducted simultaneously. Houghten *et al.*, *supra*, at 5134.

Epitope-bearing peptides and polypeptides of the invention are used to induce antibodies according to methods well known in the art. See, for instance, Sutcliffe *et al.*, *supra*; Wilson *et al.*, *supra*; Chow, M. *et al.*, *Proc. Natl. Acad. Sci. (USA) 82*:910-914; and Bittle, F.J. *et al.*, *J. Gen. Virol. 66*:2347-2354 (1985). Generally, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling of the peptide to a macromolecular carrier, such as keyhole limpet hemacyanin (KLH) or tetanus toxoid. For instance, peptides containing cysteine may be coupled to carrier using a linker such as m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carrier using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice are immunized with either free or carrier-coupled peptides, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 μg peptide or carrier protein and Freund's adjuvant. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody which can be detected, for example, by ELISA assay using free peptide adsorbed to a solid surface. The titer of anti-peptide antibodies in serum from an immunized animal may be increased by selection of anti-peptide antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

Immunogenic epitope-bearing peptides of the invention, i.e., those parts of a protein that elicit an antibody response when the whole protein is the immunogen, are identified according to methods known in the art. For instance, Geysen *et al.*, 1984, *supra*, discloses a procedure for rapid concurrent synthesis on solid supports of hundreds of peptides of sufficient purity to react in an enzyme-linked immunosorbent assay. Interaction of synthesized peptides with antibodies is then easily detected without removing them from the support. In this manner a peptide bearing an immunogenic epitope of a desired protein may be identified routinely by one of ordinary skill in the art. For instance, the immunologically important epitope in the coat protein of foot-and-mouth disease virus was located by Geysen *et al.* with a resolution of seven amino acids by synthesis of an overlapping set of all 208 possible hexapeptides covering the entire 213 amino acid sequence of the protein. Then, a complete replacement set of peptides in which all 20 amino acids were substituted in turn at every position within the epitope were synthesized, and the particular amino acids conferring specificity for the reaction with antibody were determined. Thus, peptide analogs of the epitope-bearing peptides of the invention can be made routinely by this method. U.S. Patent No. 4,708,781 to Geysen (1987) further describes this method of identifying a peptide bearing an immunogenic epitope of a desired protein.

Further still, U.S. Patent No. 5,194,392 to Geysen (1990) describes a general method of detecting or determining the sequence of monomers (amino acids or other compounds) which is a topological equivalent of the epitope (i.e., a "mimotope") which is complementary to a particular paratope (antigen binding site) of an antibody of interest. More generally, U.S. Patent No. 4,433,092 to Geysen (1989) describes a method of detecting or determining a sequence of monomers which is a topographical equivalent of a ligand which is complementary to the ligand binding site of a particular receptor of interest. Similarly, U.S. Patent No. 5,480,971 to Houghten, R.A. *et al.* (1996) on Peralkylated Oligopeptide Mixtures discloses linear C_1-C_7-alkyl peralkylated oligopeptides and sets and libraries of such peptides, as well as methods for using such oligopeptide sets and

libraries for determining the sequence of a peralkylated oligopeptide that preferentially binds to an acceptor molecule of interest. Thus, non-peptide analogs of the epitope-bearing peptides of the invention also can be made routinely by these methods.

The entire disclosure of each document cited in this section on "Polypeptides and Peptides" is hereby incorporated herein by reference

As one of skill in the art will appreciate, chitinase alpha and chitinase alpha-2 polypeptides of the present invention and the epitope-bearing fragments thereof described above can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life *in vivo*. This has been shown, e.g., for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins (EPA 394,827; Traunecker *et al., Nature 331*:84- 86 (1988)). Fusion proteins that have a disulfide-linked dimeric structure due to the IgG part can also be more efficient in binding and neutralizing other molecules than the monomeric protein or protein fragment alone (Fountoulakis *et al., J Biochem 270*:3958-3964 (1995)).

Chromosome Assays

The nucleic acid molecules of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease..

In certain preferred embodiments in this regard, the cDNA herein disclosed is used to clone genomic DNA of a chitinase alpha protein gene. This can be accomplished using a variety of well known techniques and libraries, which generally are available commercially. The genomic DNA then is used for *in situ* chromosome mapping using well known techniques for this purpose. Typically, in accordance with routine procedures for chromosome mapping, some trial and error may be necessary to identify a genomic probe that gives a good *in situ* hybridization signal.

In some cases, in addition, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3′ untranslated region of the gene is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified portion.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of portions from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include in situ hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence *in situ* hybridization ("FISH") of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with probes from the cDNA as short as 50 or 60 bp. For a review of this technique, see Verma *et al.*, *Human Chromosomes: A Manual of Basic Techniques*, Pergamon Press, New York (1988).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, *Mendelian Inheritance in Man*, available on-line through Johns Hopkins University, Welch Medical Library. The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region associated with the disease could be one of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb).

Identification of Receptor for the Chitinase Alpha or Chitinase Alpha-2 Ligand

This invention further provides a method for identification of the receptor for the chitinase alpha or chitinase alpha-2 ligand. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting (Coligan, *et al.*, *Current Protocols in Immun. 1*(2), Chapter 5, (1991)). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the chitinase alpha or chitinase alpha-2 ligand, and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the chitinase alpha or chitinase alpha-2 ligand. Transfected cells which are grown on glass slides are exposed to labeled chitinase alpha or chitinase alpha-2 ligand, which can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation

and incubation, the slides are subjected to auto-radiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an iterative sub-pooling and re-screening process, eventually yielding a single clone that encodes the putative receptor.

As an alternative approach for receptor identification, labeled ligand can be photoaffinity linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the ligand-receptor can be excised, resolved into peptide fragments, and subjected to protein microsequencing. The amino acid sequence obtained from microsequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

The method for determining whether a ligand can bind to the chitinase alpha or chitinase alpha-2 receptor comprises transfecting a cell population (one presumed not to contain the receptor) with the appropriate vector expressing the chitinase alpha or chitinase alpha-2 receptor, such that the cell will now express the receptor. A suitable response system is obtained by transfection of the DNA into a suitable host containing the desired second messenger pathways including cAMP, ion channels, phosphoinositide kinase, or calcium response. Such a transfection system provides a response system to analyze the activity of various ligands exposed to the cell. Ligands chosen could be identified through a rational approach by taking known ligands that interact with similar types of receptors or using small molecules, cell supernatants or extracts or natural products.

The present invention further provides a method of screening drugs to identify those which block (antagonists) interaction of ligand to receptor. A chitinase alpha or chitinase alpha-2 antagonist is a compound which eliminates or reduces the natural biological functions of chitinase alpha or chitinase alpha-2. As an example, a mammalian cell or membrane preparation expressing the chitinase alpha or chitinase alpha-2 receptor would be incubated with labeled ligand in the presence of the drug. The ability of the drug to block this interaction could then be measured. Alternatively, the response of a known second messenger system

following interaction of ligand and receptor would be measured and compared in the presence or absence of the drug. Such second messenger systems include but are not limited to, cAMP guanylate cyclase, ion channels or phosphoinositide hydrolysis.

Potential antagonists include an antibody, or in some cases, an oligopeptide, which binds to the polypeptide. Alternatively, a potential antagonist may be a closely related protein which binds to the receptor sites, however, they are inactive forms of the polypeptide and thereby prevent the action of chitinase alpha or chitinase alpha-2 since receptor sites are occupied.

Chitinase alpha or chitinase alpha-2 can also be used in the production of antibodies. The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the polypeptides corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptides itself. In this manner, even a sequence encoding only a fragment of the polypeptides can be used to generate antibodies binding the whole native polypeptides. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, *Nature 256*:495-497 (1975)), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today 4*:72 (1983)), and the EBV-hybridoma technique to

produce human monoclonal antibodies (Cole*et al.*, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985), pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic chitinase alpha or chitinase alpha-2 products of this invention. Also, transgenic mice may be used to express humanized antibodies to immunogenic chitinase alpha or chitinase alpha-2 products of this invention.

Another potential antagonist is an antisense construct prepared using antisense technology. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes for the mature polypeptides of the present invention, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix-see Lee *et al.*, *Nucl. Acids Res.* 6:3073 (1979); Cooney *et al.*, *Science 241*:456 (1988); and Dervan *et al.*, *Science 251*:1360 (1991)), thereby preventing transcription and the production of chitinase alpha. The antisense RNA oligonucleotide hybridizes to the mRNA in vivo and blocks translation of the mRNA molecule into chitinase alpha or chitincase alpha-2 polypeptide (Antisense - Okano, *J. Neurochem.* 56:560 (1991); *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*, CRC Press, Boca Raton, FL (1988)). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of chitinase alpha or chitinase alpha-2.

Potential antagonists also include a small molecule which binds to and occupies the catalytic site of the polypeptide thereby making the catalytic site inaccessible to substrate such that normal biological activity is prevented. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules.

Therapeutic Applications of Chitinase Alpha, Chitinase Alpha-2 and Antagonists Thereof

Treatment of pathological conditions characterized by elevated plasma levels of chitinase alpha or chitinase alpha-2

The antagonists, described above, may be employed to treat rheumatoid arthritis, atherosclerosis, inflammation, or other tissue remodeling disease processes, that may be characterized by elevated plasma levels of chitinase alpha and/or chitinase alpha-2. The antagonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as hereinafter described.

It is believed that human chitotriosidase is involved in tissue remodeling. Human chitotriosidase is homologous to microbial chitinases. In particular, the protein shares a similar cysteine motif and homology to the active site of microbial chitinases. Accordingly, chitotriosidase can be used in the development of treatments of tissue remodeling diseases and in the diagnosis of these diseases. Similarly, human chitinase alpha or chitinase alpha-2 can be used the same way.

It is believed that human serum chitotriosidase is an extracellular matrix remodeling factor involved in the remodeling process occurring in diseases. Activity against chitotriosidase substrates is indicative of a function in degrading extracellular matrix substrates with a similar carbohydrate structure to chitin. Chitotriosidase was found during morphological differentiation of monocytes into macrophages. These macrophages can be associated with different sites of inflammation that occur in various disease states, including but not limited to atherosclerosis, rheumatoid arthritis, acute inflammatory disease, and inflammation due to injury.

Accordingly, it is believed that human serum chitotriosidase, and now chitinase alpha and chitinase alpha-2, are useful targets for the inhibition of tissue remodeling disease processes. Purified human chitinase alpha or chitinase alpha-2 protein could thus be tested for various *in vitro* assays for tissue remodeling to determine the extent to which this protein promotes these processes. Therapeutics would be defined as molecules that specifically inhibit chitinase alpha or chitinase

alpha-2 function as it relates to tissue remodeling activity in these assays and includes neutralizing antibodies as described above. The invention also relates to methods of treating a patient with a tissue remodeling disorder by administering a therapeutically effective amount of a chitinase alpha or chitinase alpha-2 antagonist or antibody.

Treatment of fungal infection

In addition, the invention relates to methods of treating fungal infection by administering a therapeutically effective amount of chitinase alpha or chitinase alpha-2 to a patient. Chitinases expressed in human tissues are expected to confer protection against fungi since chitin is a component of the fungal cell wall, and chitinase cleaves this polymer. Preferably, the chitinase alpha variant having an amino acid sequence shown in Figure 2 (SEQ ID NO: 4) is used for treating fungal infection.

Modes of administration

Given the activities modulated by chitinase alpha and chitinase alpha-2, it is readily apparent that a substantially altered (increased or decreased) level of expression of these proteins in an individual compared to the standard or "normal" level produces pathological conditions such as those described above. It will also be appreciated by one of ordinary skill that, since the chitinase alpha and chitinase alpha-2 proteins of the invention are translated with a leader peptide suitable for secretion of the mature protein from cells, when chitinase alpha or chitinase alpha-2 protein (particularly the mature form) is added from an exogenous source to cells, tissues or the body of an individual, the protein will exert its modulating activities on any of its target cells. Therefore, it will be appreciated that conditions caused by a decrease in the standard or normal level of chitinase alpha or chitinase alpha-2 activity in an individual, can be treated by administration of these proteins. Thus, the invention further provides a method of treating an individual in need of

an increased level of chitinase alpha or chitinase alpha-2 activity comprising administering to such an individual a pharmaceutical composition comprising an amount of an isolated chitinase alpha and/or chitinase alpha-2 polypeptide of the invention, particularly the mature form, effective to increase the chitinase alpha and/or chitinase alpha-2 activity level in such an individual.

The chitninase alpha and/or chitinase polypeptide composition to be used in the therapy will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with chitinase alpha polypeptide alone), the site of delivery of the composition, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" of chitinase alpha or chitinase alpha-2 polypeptide for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of chitinase alpha or chitinase alpha-2 polypeptide administered parenterally per dose will be in the range of about 1 μg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the chitinase alpha or chitinase alpha-2 polypeptide is typically administered at a dose rate of about 1 μg/kg/hour to about 50 μg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed.

The chitinase alpha or chitinase alpha-2 polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. *et al.*, *Biopolymers 22*:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (R. Langer *et al.*, *J. Biomed. Mater. Res. 15*:167-277 (1981), and R. Langer, *Chem. Tech. 12*:98-105 (1982)), ethylene

vinyl acetate (R. Langer *et al., Id.*) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release chitinase alpha or chitinase alpha-2 polypeptide compositions also include liposomally entrapped polypeptides. Liposomes containing chitinase alpha or chitinase alpha-2 polypeptide are prepared by methods known *per se*: DE 3,218,121; Epstein *et al., Proc. Natl. Acad. Sci. (USA) 82*:3688-3692 (1985); Hwang *et al., Proc. Natl. Acad. Sci. (USA) 77*:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal chitinase alpha or chitinase alpha-2 polypeptide therapy.

For parenteral administration, in one embodiment, the chitinase alpha or chitinase alpha-2 polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids

or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as s rum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The chitinase alpha and chitinase alpha-2 polypeptides are typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Polypeptides to be used for therapeutic administration must be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic chitinase alpha and chitianse alpha-2 polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Chitinase alpha and chitinase alpha-2 polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous chitinase alpha polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized chitinase alpha polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval

by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 1

Bacterial Expression and Purification of Chitinase Alpha

The DNA sequence encoding chitinase alpha, ATCC No. 97529, is initially amplified using PCR oligonucleotide primers corresponding to the 5' sequences of the processed chitinase alpha protein (minus the signal peptide sequence) and the vector sequences 3' to the gene. Additional nucleotides containing restriction sites to facilitate cloning are added to the 5' and 3' sequences respectively.

The 5' oligonucleotide primer has the sequence CGCGGATCCATGGGAGCAACCACCATGGACCAG (SEQ ID NO: 5) and contains a BamHI restriction site followed by 24 nucleotides of chitinase alpha coding sequence starting from the presumed terminal amino acid of the processed protein codon.

The 3' primer sequence CGCGGATCCGCTTCTCTGTAAGTTAATCCT (SEQ ID NO:6) contains complementary sequences to the BamHI site and is followed by 21 nucleotides of chitinase-alpha DNA insert. The restriction sites correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc. 9259 Eton Avenue, Chatsworth, CA, 91311). pQE-9 encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter operator (P/O), a ribosome binding site (RBS), a 6-His tag and restriction enzyme sites. pQE-9 is then digested with BamHI.

The amplified sequences are ligated into pQE-9 and inserted in frame with the sequence encoding for the histidine tag and the RBS. The ligation mixture is then used to transform the *E. coli* strain available from Qiagen under the trademark M15/rep 4 by the procedure described in Sambrook, J. *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Laboratory Press,

(1989). M15/rep4 contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies were selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 µg/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells were grown to an optical density 600 ($O.D._{600}$) of between 0.4 and 0.6. IPTG ("Isopropyl-B-D-thiogalacto pyranoside") is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression. Cells are grown an extra 3 to 4 hours. Cells were then harvested by centrifugation. The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl. After clarification, solubilized chitinase alpha is purified from this solution by chromatography on a Nickel-Chelate column under conditions that allow for tight binding by proteins containing the 6-His tag (Hochuli, E. *et al.*, *J. Chromatography 411*:177-184 (1984)). Chitinase-alpha (>90% pure) is eluted from the column in 6 molar guanidine HCl pH 5.0 and for the purpose of renaturation adjusted to 3 molar guanidine HCl, 100mM sodium phosphate, 10 mmolar glutathione (reduced) and 2 mmolar glutathione (oxidized). After incubation in this solution for 12 hours, the protein is dialyzed to 10 mmolar sodium phosphate.

Example 2

Cloning and Expression of Chitinase Alpha in Baculovirus

The DNA sequence encoding the full length chitinase alpha protein, ATCC No. 97529, is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene:

The 5' primer has the sequence

CGCGGATCCGCCATCATGGGAGCAACCACCATG (SEQ ID NO: 7) and contains a BamHI restriction site followed by 6 nucleotides resembling an efficient signal for the initiation of translation in eukaryotic cells (Kozak, M., *J. Mol. Biol. 196*:947-950 (1987) which is just behind the first 280 nucleotides of the chitinase-alpha gene (the initiation codon for translation "ATG" is underlined).

The 3' primer has the sequence CGCGGATCCGCTTCTCTGTAAGTTAATCCT (SEQ ID NO: 8) and contains the cleavage site for the restriction endonuclease BamHI and 21 nucleotides complementary to the 3' non-translated sequence of the chitinase-alpha gene. The amplified sequences are isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment is then digested with the endonuclease BamHI, and then purified again on a 1% agarose gel. This fragment is designated F2.

The vector pA2 (modification of pVL941 vector, discussed below) is used for the expression of the chitinase alpha protein using the baculovirus expression system (for review see: Summers, M.D. and Smith, G.E. *A manual of methods for baculovirus vectors and insect cell culture procedures*, Texas Agricultural Experimental Station Bulletin No. 1555 (1987)). This expression vector contains the strong polyhedrin promoter of the Autographa californica nuclear polyhedrosis virus (AcMNPV) followed by the recognition sites for the restriction endonuclease BamHI. The polyadenylation site of the simian virus (SV)40 is used for efficient polyadenylation. For an easy selection of recombinant virus the beta-galactosidase gene from E.coli is inserted in the same orientation as the polyhedrin promoter followed by the polyadenylation signal of the polyhedrin gene. The polyhedrin sequences are flanked at both sides by viral sequences for the cell-mediated homologous recombination of co-transfected wild-type viral DNA. Many other baculovirus vectors could be used in place of pRG1 such as pAc373, pVL941 and pAcIM1 (Luckow, V.A. and Summers, M.D., *Virology 170*:31-39).

The plasmid is digested with the restriction enzyme BamHI and then dephosphorylated using calf intestinal phosphatase by procedures known in the art. The DNA is then isolated from a 1% agarose gel using the commercially

available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.). This vector DNA is designated V2.

Fragment F2 and the dephosphorylated plasmid V2 is ligated with T4 DNA ligase. *E.coli* HB101 cells were then transformed and bacteria identified that contained the plasmid (pBac chitinase alpha) with the chitinase-alpha gene using the enzyme BamHI. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five μg of the plasmid pBacchitinase alpha is co-transfected with 1.0 μg of a commercially available linearized baculovirus ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA.) using the lipofection method (Felgner *et al.*, *Proc. Natl. Acad. Sci. (USA) 84*:7413-7417 (1987)).

Oneμg of BaculoGold™ virus DNA and 5 μg of the plasmid pBac chitinase alpha are mixed in a sterile well of a microtiter plate containing 50 μl of serum free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to the Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is rocked back and forth to mix the newly added solution. The plate is then incubated for 5 hours at 27°C. After 5 hours, the transfection solution is removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. The plate is put back into an incubator and cultivation continued at 27°C for four days.

After four days, the supernatant is collected and a plaque assay is performed similar to the assay described by Summers and Smith (supra). As a modification, an agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used, which allows an easy isolation of blue stained plaques. (A detailed description of a "plaque assay" can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10).

Four days after the serial dilution, the virus is added to the cells, and blue stained plaques are picked with the tip of an Eppendorf pipette. The agar containing the recombinant viruses is then resuspended in an Eppendorf tube containing 200 μl of Grace's medium. The agar is removed by a brief centrifugation and the supernatant containing the recombinant baculovirus was used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then stored at 4°C.

Sf9 cells were grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells were infected with the recombinant baculovirus V-chitinase-alpha at a multiplicity of infection (MOI) of 2. Six hours later, the medium is removed and replaced with SF900 II medium minus methionine and cysteine (Life Technologies Inc., Gaithersburg). 42 hours later 5 μCi of ^{35}S-methionine and 5 μCi ^{35}S cysteine (Amersham) are added. The cells are further incubated for 16 hours before they are harvested by centrifugation and the labelled proteins visualized by SDS-PAGE and autoradiography.

Example 3

Cloning and Expression in Mammalian Cells

Most of the vectors used for the transient expression of the chitinase alpha protein gene sequence in mammalian cells should carry the SV40 origin of replication. This allows the replication of the vector to high copy numbers in cells (e.g. COS cells) which express the T antigen required for the initiation of viral DNA synthesis. Any other mammalian cell line can also be utilized for this purpose.

A typical mammalian expression vector contains the promoter element, which mediates the initiation of transcription of mRNA, the protein coding sequence, and signals required for the termination of trancription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription can be achieved with the early and late

promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g. RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular signals can also be used (e.g. human actin promoter). Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) and pBC12MI (ATCC 67109). Mammalian host cells that could be used include, human Hela, 283, H9 and Jurkart cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, African green monkey cells, quail QC1-3 cells, mouse L cells and Chinese hamster ovary cells.

Alternatively, the gene can be expressed in stable cell lines that contain the gene integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) is a useful marker to develop cell lines that carry several hundred or even several thousand copies of the gene of interest. Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy *et al., Biochem J. 227:*277-279 (1991); Bebbington *et al., Bio/Technology 10:*169-175 (1992)). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) cells are often used for the production of proteins.

The expression vectors pC1 and pC4 contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen *et al., Molecular and Cellular Biology,* 438-4470 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart *et al., Cell 41:*521-530 (1985)). Multiple cloning sites, e.g. with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors contain in addition the 3′ intron, the polyadenylation and termination signal of the rat preproinsulin gene.

Example 3(a) Expression in COS Cells

The expression plasmid, chitinase alpha HA, is made by cloning a cDNA encoding chitinase alpha into the expression vector pcDNAI/Amp (which can be obtained from Invitrogen, Inc.).

The expression vector pcDNAI/amp contains: (1) an *E.coli* origin of replication effective for propagation in *E. coli* and other prokaryotic cells; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin of replication for propagation in eukaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron, and a polyadenylation signal arranged so that a cDNA conveniently can be placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker.

A DNA fragment encoding the entire chitinase alpha precursor and an HA tag fused in frame to its 3' end is cloned into the polylinker region of the vector so that recombinant protein expression is directed by the CMV promoter. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein described by Wilson *et al.*, *Cell 37*:767 (1984). The fusion of the HA tag to the target protein allows easy detection of the recombinant protein with an antibody that recognizes the HA epitope.

The plasmid construction strategy is as follows.

The chitinase alpha cDNA of the deposited clone is amplified using primers that contain convenient restriction sites, much as described above regarding the construction of expression vectors for expression of chitinase alpha in *E. coli*. To facilitate detection, purification and characterization of the expressed chitinase alpha, one of the primers contains a hemagglutinin tag ("HA tag") as described above.

Suitable primers include that following, which are used in this example. The 5' primer, containing the underlined BamHI site, an AUG start codon and 6 codons of the 5' coding region has the following sequence:

5' CGC GGA TCC GCC ATC ATG GGA GCA ACC ACC ATG (SEQ ID NO: 9)

The 3' primer, containing complementary sequences to the BamHI site, a stop codon, 9 codons thereafter forming the hemagglutinin HA tag, and 18 bp of 3' coding sequence (at the 3' end) has the following sequence:

5' CGC GGA TCC TCA AGC GTA GTC TGG GAC GTC GTA TGG GTA CAG GGA GCC AAG GCT TCT (SEQ ID NO: 10)

The PCR amplified DNA fragment and the vector, pcDNAI/Amp, are digested with BamHI and then ligated. The ligation mixture is transformed into *E. coli* strain SURE (available from Stratagene Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA 92037), and the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis and gel sizing for the presence of the chemokine alpha-encoding fragment.

For expression of recombinant chemokine alpha, COS cells are transfected with an expression vector, as described above, using DEAE-DEXTRAN, as described, for instance, in Sambrook, J. *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Laboratory Press, Cold Spring Harbor, New York (1989). Cells are incubated under conditions for expression of chemokine alpha by the vector.

Expression of the chemokine alpha HA fusion protein is detected by radiolabelling and immunoprecipitation, using methods described in, for example Harlow *et al.*, *Antibodies: A Laboratory Manual*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988). To this end, two days after transfection, the cells are labeled by incubation in media containing ^{35}S-cysteine for 8 hours. The cells and the media are collected, and the cells are washed and the lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson *et al.* cited above. Proteins are precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The

precipitated proteins then are analyzed by SDS-PAGE gels and autoradiography. An expression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

Example 3(b): Cloning and Expression in CHO Cells

The DNA sequence encoding chitinase alpha protein is amplified using PCR oligonucleotide primers specific to the amino terminal sequence of the protein and to vector sequences 3′ to the gene. Additional nucleotides containing restriction sites to facilitate cloning are added to the 5′ and 3′ sequences respectively. The restrictions sites are convenient to restriction enzyme sites in the CHO expression vector PC4.

The amplified chitinase alpha protein DNA and the vector PC4 both are digested with BamHI and the digested DNAs then ligated together. Insertion of the chitinase alpha protein DNA into the BamHI restricted vector places the protein coding region downstream of and operably linked to the vector's promoter. The ligation mixture is transformed into competent CHO cells using standard procedures as described, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

Example 4

Expression Pattern of Chitinase Alpha in Human Tissue

Northern blot analysis is carried out to examine chitinae alpha gene expression in human tissues, using methods described by, among others, Sambrook *et al.*, cited above. A cDNA probe containing the entire nucleotide sequence of the chitinase alpha protein (SEQ ID NO:1) is labeled with ^{32}P using the *redi*prime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labelling, the probe is purified using a CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labelled probe is then used to examine various human tissues for chitinase alpha mRNA.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) are obtained from Clontech and are examined with labelled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and films developed according to standard procedures.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, within the scope of the appended claims, the invention may be practiced otherwise than as particularly described.

The disclosures of all patents, patent applications and publications referred to herein are hereby incorporated by reference.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: HUMAN GENOME SCIENCES, INC.
9410 KEY WEST AVENUE
ROCKVILLE, MD 20850
APPLICANT/INVENTORS: THOTAKURA, NAGESWARA RAO
CHOPRA, ARVIND
CHOI, GIL
GENTZ, REINER
ROSEN, CRAIG

(ii) TITLE OF INVENTION: CHITINASE

(iii) NUMBER OF SEQUENCES: 12

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
(B) STREET: 1100 NEW YORK AVENUE, SUITE 600
(C) CITY: WASHINGTON
(D) STATE: D.C.
(E) COUNTRY: US
(F) ZIP: 20005-3934

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: TBA
(B) FILING DATE: 09-AUG-96
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: GOLDSTEIN, JORGE A.
(B) REGISTRATION NUMBER: 29,021
(C) REFERENCE/DOCKET NUMBER: 1488.046PC00

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 202-371-2600
(B) TELEFAX: 202-371-2540

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1595 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 82..1344

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GAATTCGGCA CGAGGGCTGT CGAAACCTCA GTGGATAAAA GACCTAGAGA ATGTGTATCC       60

CAGAAGAAGC TGGCCAAGGA T ATG GGA GCA ACC ACC ATG GAC CAG AAG TCT         111
                        Met Gly Ala Thr Thr Met Asp Gln Lys Ser
                          1               5                  10

CTC TGG GCA GGT GTA GTG GTC TTG CTG CTT CTC CAG GGA GAG ATG GGG         159
Leu Trp Ala Gly Val Val Val Leu Leu Leu Leu Gln Gly Glu Met Gly
                 15                  20                  25

TTT TGC TAT GTT GCC AGA GCT GGT CTT GAA CTC CTG GGC TCA AGA AGT         207
Phe Cys Tyr Val Ala Arg Ala Gly Leu Glu Leu Leu Gly Ser Arg Ser
             30                  35                  40

CCT CCT GCC TCA GCC TCC CAA AGT GCT GGG ATA ACA GGA TCT GCC TAC         255
Pro Pro Ala Ser Ala Ser Gln Ser Ala Gly Ile Thr Gly Ser Ala Tyr
         45                  50                  55

AAA CTG GTT TGC TAC TTT ACC AAC TGG TCC CAG GAC CGG CAG GAA CCA         303
Lys Leu Val Cys Tyr Phe Thr Asn Trp Ser Gln Asp Arg Gln Glu Pro
     60                  65                  70

GGA AAA TTC ACC CCT GAG AAT ATT GAC CCC TTC CTA TGC TCT CAT CTC         351
Gly Lys Phe Thr Pro Glu Asn Ile Asp Pro Phe Leu Cys Ser His Leu
 75                  80                  85                  90

ATC TAT TCA TTC GCC AGC ATC GAA AAC AAC AAG GTT ATC ATC AAG GAC         399
Ile Tyr Ser Phe Ala Ser Ile Glu Asn Asn Lys Val Ile Ile Lys Asp
                 95                 100                 105

AAG AGT GAA GTG ATG CTC TAC CAG ACC ATC AAC AGT CTC AAA ACC AAG         447
Lys Ser Glu Val Met Leu Tyr Gln Thr Ile Asn Ser Leu Lys Thr Lys
            110                 115                 120

AAT CCC AAA CTG AAA ATT CTC TTG TCC ATT GGA GGG TAC CTG TTT GGT         495
Asn Pro Lys Leu Lys Ile Leu Leu Ser Ile Gly Gly Tyr Leu Phe Gly
        125                 130                 135

TCC AAA GGG TTC CAC CCT ATG GTG GAT TCT TCT ACA TCA CGC TTG GAA         543
```

```
Ser Lys Gly Phe His Pro Met Val Asp Ser Ser Thr Ser Arg Leu Glu
    140                 145                 150

TTC ATT AAC TCC ATA ATC CTG TTT CTG AGG AAC CAT AAC TTT GAT GGA      591
Phe Ile Asn Ser Ile Ile Leu Phe Leu Arg Asn His Asn Phe Asp Gly
155                 160                 165                 170

CTG GAT GTA AGC TGG ATC TAC CCA GAT CAG AAA GAA AAC ACT CAT TTC      639
Leu Asp Val Ser Trp Ile Tyr Pro Asp Gln Lys Glu Asn Thr His Phe
                175                 180                 185

ACT GTG CTG ATT CAT GAG TTA GCA GAA GCC TTT CAG AAG GAC TTC ACA      687
Thr Val Leu Ile His Glu Leu Ala Glu Ala Phe Gln Lys Asp Phe Thr
            190                 195                 200

AAA TCC ACC AAG GAA AGG CTT CTC TTG ACT GCG GGC GTA TCT GCA GGG      735
Lys Ser Thr Lys Glu Arg Leu Leu Leu Thr Ala Gly Val Ser Ala Gly
        205                 210                 215

AGG CAA ATG ATT GAT AAC AGC TAT CAA GTT GAG AAA CTG GCA AAA GAT      783
Arg Gln Met Ile Asp Asn Ser Tyr Gln Val Glu Lys Leu Ala Lys Asp
    220                 225                 230

CTG GAT TTC ATC AAC CTC CTG TCC TTT GAC TTC CAT GGG TCT TGG GAA      831
Leu Asp Phe Ile Asn Leu Leu Ser Phe Asp Phe His Gly Ser Trp Glu
235                 240                 245                 250

AAG CCC CTT ATC ACT GGC CAC AAC AGC CCT CTG AGC AAG GGG TGG CAG      879
Lys Pro Leu Ile Thr Gly His Asn Ser Pro Leu Ser Lys Gly Trp Gln
                255                 260                 265

GAC AGA GGG CCA AGC TCC TAC TAC AAT GTG GAA TAT GCT GTG GGG TAC      927
Asp Arg Gly Pro Ser Ser Tyr Tyr Asn Val Glu Tyr Ala Val Gly Tyr
            270                 275                 280

TGG ATA CAT AAG GGA ATG CCA TCA GAG AAG GTG GTC ATG GGC ATC CCC      975
Trp Ile His Lys Gly Met Pro Ser Glu Lys Val Val Met Gly Ile Pro
        285                 290                 295

ACA TAT GGG CAC TCC TTC ACA CTG GCC TCT GCA GAA ACC ACC GTG GGG     1023
Thr Tyr Gly His Ser Phe Thr Leu Ala Ser Ala Glu Thr Thr Val Gly
    300                 305                 310

GCC CCT GCC TCT GGC CCT GGA GCT GCT GGA CCC ATC ACA GAG TCT TCA     1071
Ala Pro Ala Ser Gly Pro Gly Ala Ala Gly Pro Ile Thr Glu Ser Ser
315                 320                 325                 330

GGC TTC CTG GCC TAT TAT GAG ATC TGC CAG TTC CTG AAA GGA GCC AAG     1119
Gly Phe Leu Ala Tyr Tyr Glu Ile Cys Gln Phe Leu Lys Gly Ala Lys
                335                 340                 345

ATC ACG AGG CTC CAG GAT CAG CAG GTT CCC TAC GCA GTC AAG GGG AAC     1167
Ile Thr Arg Leu Gln Asp Gln Gln Val Pro Tyr Ala Val Lys Gly Asn
            350                 355                 360
```

```
CAG TGG GTG GGC TAT GAT GAT GTG AAG AGT ATG GAG ACC AAG GTT CAG      1215
Gln Trp Val Gly Tyr Asp Asp Val Lys Ser Met Glu Thr Lys Val Gln
        365                 370                 375

TTC TTA AAG AAT TTA AAC CTG GGA GGA GCC ATG ATC TGG TCT ATT GAC      1263
Phe Leu Lys Asn Leu Asn Leu Gly Gly Ala Met Ile Trp Ser Ile Asp
    380                 385                 390

ATG GAT GAC TTC ACT GGC AAA TCC TGC AAC CAG GGC CCT TAC CCT CTT      1311
Met Asp Asp Phe Thr Gly Lys Ser Cys Asn Gln Gly Pro Tyr Pro Leu
395                 400                 405                 410

GTC CAA GCA GTC AAG AGA AGC CTT GGC TCC CTG TGAAGGATTA ACTTACAGAG    1364
Val Gln Ala Val Lys Arg Ser Leu Gly Ser Leu
                415                 420

AAGCAGGCAA GATGACCTTG CTGCCTGGGG CCTGCTCTCT CCCAGGAATT CTCATGTGGG    1424

ATTCCCCTTG CCAGGCCGGC CTTTGGATCT CTCTTCCAAG CCTTTCCTGA CTTCCTCTTA    1484

GATCATAGAT TGGACCTGGT TTTGTTTTCC TGCAGCTGAT GACTTGTTGC CCTGAAGTAC    1544

AATAAAAAAA ATTCATTTTG CTCCAGTAAA AAAAAAAAAA AAAAACTCGA G              1595
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 421 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Gly Ala Thr Thr Met Asp Gln Lys Ser Leu Trp Ala Gly Val Val
 1               5                  10                  15

Val Leu Leu Leu Leu Gln Gly Glu Met Gly Phe Cys Tyr Val Ala Arg
            20                  25                  30

Ala Gly Leu Glu Leu Leu Gly Ser Arg Ser Pro Pro Ala Ser Ala Ser
        35                  40                  45

Gln Ser Ala Gly Ile Thr Gly Ser Ala Tyr Lys Leu Val Cys Tyr Phe
    50                  55                  60

Thr Asn Trp Ser Gln Asp Arg Gln Glu Pro Gly Lys Phe Thr Pro Glu
65                  70                  75                  80

Asn Ile Asp Pro Phe Leu Cys Ser His Leu Ile Tyr Ser Phe Ala Ser
                85                  90                  95

Ile Glu Asn Asn Lys Val Ile Ile Lys Asp Lys Ser Glu Val Met Leu
```

100 105 110

Tyr Gln Thr Ile Asn Ser Leu Lys Thr Lys Asn Pro Lys Leu Lys Ile
115 120 125

Leu Leu Ser Ile Gly Gly Tyr Leu Phe Gly Ser Lys Gly Phe His Pro
130 135 140

Met Val Asp Ser Ser Thr Ser Arg Leu Glu Phe Ile Asn Ser Ile Ile
145 150 155 160

Leu Phe Leu Arg Asn His Asn Phe Asp Gly Leu Asp Val Ser Trp Ile
165 170 175

Tyr Pro Asp Gln Lys Glu Asn Thr His Phe Thr Val Leu Ile His Glu
180 185 190

Leu Ala Glu Ala Phe Gln Lys Asp Phe Thr Lys Ser Thr Lys Glu Arg
195 200 205

Leu Leu Leu Thr Ala Gly Val Ser Ala Gly Arg Gln Met Ile Asp Asn
210 215 220

Ser Tyr Gln Val Glu Lys Leu Ala Lys Asp Leu Asp Phe Ile Asn Leu
225 230 235 240

Leu Ser Phe Asp Phe His Gly Ser Trp Glu Lys Pro Leu Ile Thr Gly
245 250 255

His Asn Ser Pro Leu Ser Lys Gly Trp Gln Asp Arg Gly Pro Ser Ser
260 265 270

Tyr Tyr Asn Val Glu Tyr Ala Val Gly Tyr Trp Ile His Lys Gly Met
275 280 285

Pro Ser Glu Lys Val Val Met Gly Ile Pro Thr Tyr Gly His Ser Phe
290 295 300

Thr Leu Ala Ser Ala Glu Thr Thr Val Gly Ala Pro Ala Ser Gly Pro
305 310 315 320

Gly Ala Ala Gly Pro Ile Thr Glu Ser Ser Gly Phe Leu Ala Tyr Tyr
325 330 335

Glu Ile Cys Gln Phe Leu Lys Gly Ala Lys Ile Thr Arg Leu Gln Asp
340 345 350

Gln Gln Val Pro Tyr Ala Val Lys Gly Asn Gln Trp Val Gly Tyr Asp
355 360 365

Asp Val Lys Ser Met Glu Thr Lys Val Gln Phe Leu Lys Asn Leu Asn
370 375 380

Leu Gly Gly Ala Met Ile Trp Ser Ile Asp Met Asp Asp Phe Thr Gly
385 390 395 400

```
Lys Ser Cys Asn Gln Gly Pro Tyr Pro Leu Val Gln Ala Val Lys Arg
                405                 410                 415

Ser Leu Gly Ser Leu
            420
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1595 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 82..1344

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GAATTCGGCA CGAGGGCTGT CGAAACCTCA GTGGATAAAA GACCTAGAGA ATGTGTATCC      60

CAGAAGAAGC TGGCCAAGGA T ATG GGA GCA ACC ACC ATG GAC CAG AAG TCT      111
                        Met Gly Ala Thr Thr Met Asp Gln Lys Ser
                          1               5                  10

CTC TGG GCA GGT GTA GTG GTC TTG CTG CTT CTC CAG GGA GAG ATG GGG      159
Leu Trp Ala Gly Val Val Val Leu Leu Leu Leu Gln Gly Glu Met Gly
                 15                  20                  25

TTT TGC TAT GTT GCC AGA GCT GGT CTT GAA CTC CTG GGC TCA AGA AGT      207
Phe Cys Tyr Val Ala Arg Ala Gly Leu Glu Leu Leu Gly Ser Arg Ser
             30                  35                  40

CCT CCT GCC TCA GCC TCC CAA AGT GCT GGG ATA ACA GGA TCT GCC TAC      255
Pro Pro Ala Ser Ala Ser Gln Ser Ala Gly Ile Thr Gly Ser Ala Tyr
         45                  50                  55

AAA CTG GTT TGC TAC TTT ACC AAC TGG TCC CAG GAC CGG CAG GAA CCA      303
Lys Leu Val Cys Tyr Phe Thr Asn Trp Ser Gln Asp Arg Gln Glu Pro
     60                  65                  70

GGA AAA TTC ACC CCT GAG AAT ATT GAC CCC TTC CTA TGC TCT CAT CTC      351
Gly Lys Phe Thr Pro Glu Asn Ile Asp Pro Phe Leu Cys Ser His Leu
 75                  80                  85                  90

ATC TAT TCA TTC GCC AGC ATC GAA AAC AAC AAG GTT ATC ATC AAG GAC      399
Ile Tyr Ser Phe Ala Ser Ile Glu Asn Asn Lys Val Ile Ile Lys Asp
                 95                 100                 105

AAG AGT GAA GTG ATG CTC TAC CAG ACC ATC AAC AGT CTC AAA ACC AAG      447
```

```
Lys Ser Glu Val Met Leu Tyr Gln Thr Ile Asn Ser Leu Lys Thr Lys
            110                 115                 120

AAT CCC AAA CTG AAA ATT CTC TTG TCC ATT GGA GGG TAC CTG TTT GGT      495
Asn Pro Lys Leu Lys Ile Leu Leu Ser Ile Gly Gly Tyr Leu Phe Gly
        125                 130                 135

TCC AAA GGG TTC CAC CCT ATG GTG GAT TCT TCT ACA TCA CGC TTG GAA      543
Ser Lys Gly Phe His Pro Met Val Asp Ser Ser Thr Ser Arg Leu Glu
    140                 145                 150

TTC ATT AAC TCC ATA ATC CTG TTT CTG AGG AAC CAT AAC TTT GAT GGA      591
Phe Ile Asn Ser Ile Ile Leu Phe Leu Arg Asn His Asn Phe Asp Gly
155                 160                 165                 170

CTG GAT GTA AGC TGG GAA TAC CCA GAT CAG AAA GAA AAC ACT CAT TTC      639
Leu Asp Val Ser Trp Glu Tyr Pro Asp Gln Lys Glu Asn Thr His Phe
                175                 180                 185

ACT GTG CTG ATT CAT GAG TTA GCA GAA GCC TTT CAG AAG GAC TTC ACA      687
Thr Val Leu Ile His Glu Leu Ala Glu Ala Phe Gln Lys Asp Phe Thr
            190                 195                 200

AAA TCC ACC AAG GAA AGG CTT CTC TTG ACT GCG GGC GTA TCT GCA GGG      735
Lys Ser Thr Lys Glu Arg Leu Leu Leu Thr Ala Gly Val Ser Ala Gly
        205                 210                 215

AGG CAA ATG ATT GAT AAC AGC TAT CAA GTT GAG AAA CTG GCA AAA GAT      783
Arg Gln Met Ile Asp Asn Ser Tyr Gln Val Glu Lys Leu Ala Lys Asp
    220                 225                 230

CTG GAT TTC ATC AAC CTC CTG TCC TTT GAC TTC CAT GGG TCT TGG GAA      831
Leu Asp Phe Ile Asn Leu Leu Ser Phe Asp Phe His Gly Ser Trp Glu
235                 240                 245                 250

AAG CCC CTT ATC ACT GGC CAC AAC AGC CCT CTG AGC AAG GGG TGG CAG      879
Lys Pro Leu Ile Thr Gly His Asn Ser Pro Leu Ser Lys Gly Trp Gln
                255                 260                 265

GAC AGA GGG CCA AGC TCC TAC TAC AAT GTG GAA TAT GCT GTG GGG TAC      927
Asp Arg Gly Pro Ser Ser Tyr Tyr Asn Val Glu Tyr Ala Val Gly Tyr
            270                 275                 280

TGG ATA CAT AAG GGA ATG CCA TCA GAG AAG GTG GTC ATG GGC ATC CCC      975
Trp Ile His Lys Gly Met Pro Ser Glu Lys Val Val Met Gly Ile Pro
        285                 290                 295

ACA TAT GGG CAC TCC TTC ACA CTG GCC TCT GCA GAA ACC ACC GTG GGG     1023
Thr Tyr Gly His Ser Phe Thr Leu Ala Ser Ala Glu Thr Thr Val Gly
    300                 305                 310

GCC CCT GCC TCT GGC CCT GGA GCT GCT GGA CCC ATC ACA GAG TCT TCA     1071
Ala Pro Ala Ser Gly Pro Gly Ala Ala Gly Pro Ile Thr Glu Ser Ser
315                 320                 325                 330
```

```
GGC TTC CTG GCC TAT TAT GAG ATC TGC CAG TTC CTG AAA GGA GCC AAG      1119
Gly Phe Leu Ala Tyr Tyr Glu Ile Cys Gln Phe Leu Lys Gly Ala Lys
                335                 340                 345

ATC ACG AGG CTC CAG GAT CAG CAG GTT CCC TAC GCA GTC AAG GGG AAC      1167
Ile Thr Arg Leu Gln Asp Gln Gln Val Pro Tyr Ala Val Lys Gly Asn
            350                 355                 360

CAG TGG GTG GGC TAT GAT GAT GTG AAG AGT ATG GAG ACC AAG GTT CAG      1215
Gln Trp Val Gly Tyr Asp Asp Val Lys Ser Met Glu Thr Lys Val Gln
        365                 370                 375

TTC TTA AAG AAT TTA AAC CTG GGA GGA GCC ATG ATC TGG TCT ATT GAC      1263
Phe Leu Lys Asn Leu Asn Leu Gly Gly Ala Met Ile Trp Ser Ile Asp
    380                 385                 390

ATG GAT GAC TTC ACT GGC AAA TCC TGC AAC CAG GGC CCT TAC CCT CTT      1311
Met Asp Asp Phe Thr Gly Lys Ser Cys Asn Gln Gly Pro Tyr Pro Leu
395                 400                 405                 410

GTC CAA GCA GTC AAG AGA AGC CTT GGC TCC CTG TGAAGGATTA ACTTACAGAG    1364
Val Gln Ala Val Lys Arg Ser Leu Gly Ser Leu
                415                 420

AAGCAGGCAA GATGACCTTG CTGCCTGGGG CCTGCTCTCT CCCAGGAATT CTCATGTGGG    1424

ATTCCCCTTG CCAGGCCGGC CTTTGGATCT CTCTTCCAAG CCTTTCCTGA CTTCCTCTTA    1484

GATCATAGAT TGGACCTGGT TTTGTTTTCC TGCAGCTGAT GACTTGTTGC CCTGAAGTAC    1544

AATAAAAAAA ATTCATTTTG CTCCAGTAAA AAAAAAAAAA AAAAACTCGA G              1595
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 421 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Gly Ala Thr Thr Met Asp Gln Lys Ser Leu Trp Ala Gly Val Val
 1               5                  10                  15

Val Leu Leu Leu Leu Gln Gly Glu Met Gly Phe Cys Tyr Val Ala Arg
            20                  25                  30

Ala Gly Leu Glu Leu Leu Gly Ser Arg Ser Pro Pro Ala Ser Ala Ser
        35                  40                  45

Gln Ser Ala Gly Ile Thr Gly Ser Ala Tyr Lys Leu Val Cys Tyr Phe
    50                  55                  60
```

```
Thr Asn Trp Ser Gln Asp Arg Gln Glu Pro Gly Lys Phe Thr Pro Glu
 65                  70                  75                  80

Asn Ile Asp Pro Phe Leu Cys Ser His Leu Ile Tyr Ser Phe Ala Ser
                 85                  90                  95

Ile Glu Asn Asn Lys Val Ile Ile Lys Asp Lys Ser Glu Val Met Leu
            100                 105                 110

Tyr Gln Thr Ile Asn Ser Leu Lys Thr Lys Asn Pro Lys Leu Lys Ile
        115                 120                 125

Leu Leu Ser Ile Gly Gly Tyr Leu Phe Gly Ser Lys Gly Phe His Pro
    130                 135                 140

Met Val Asp Ser Ser Thr Ser Arg Leu Glu Phe Ile Asn Ser Ile Ile
145                 150                 155                 160

Leu Phe Leu Arg Asn His Asn Phe Asp Gly Leu Asp Val Ser Trp Glu
                165                 170                 175

Tyr Pro Asp Gln Lys Glu Asn Thr His Phe Thr Val Leu Ile His Glu
            180                 185                 190

Leu Ala Glu Ala Phe Gln Lys Asp Phe Thr Lys Ser Thr Lys Glu Arg
        195                 200                 205

Leu Leu Leu Thr Ala Gly Val Ser Ala Gly Arg Gln Met Ile Asp Asn
    210                 215                 220

Ser Tyr Gln Val Glu Lys Leu Ala Lys Asp Leu Asp Phe Ile Asn Leu
225                 230                 235                 240

Leu Ser Phe Asp Phe His Gly Ser Trp Glu Lys Pro Leu Ile Thr Gly
                245                 250                 255

His Asn Ser Pro Leu Ser Lys Gly Trp Gln Asp Arg Gly Pro Ser Ser
            260                 265                 270

Tyr Tyr Asn Val Glu Tyr Ala Val Gly Tyr Trp Ile His Lys Gly Met
        275                 280                 285

Pro Ser Glu Lys Val Val Met Gly Ile Pro Thr Tyr Gly His Ser Phe
    290                 295                 300

Thr Leu Ala Ser Ala Glu Thr Thr Val Gly Ala Pro Ala Ser Gly Pro
305                 310                 315                 320

Gly Ala Ala Gly Pro Ile Thr Glu Ser Ser Gly Phe Leu Ala Tyr Tyr
                325                 330                 335

Glu Ile Cys Gln Phe Leu Lys Gly Ala Lys Ile Thr Arg Leu Gln Asp
            340                 345                 350

Gln Gln Val Pro Tyr Ala Val Lys Gly Asn Gln Trp Val Gly Tyr Asp
```

355 360 365

Asp Val Lys Ser Met Glu Thr Lys Val Gln Phe Leu Lys Asn Leu Asn
370 375 380

Leu Gly Gly Ala Met Ile Trp Ser Ile Asp Met Asp Asp Phe Thr Gly
385 390 395 400

Lys Ser Cys Asn Gln Gly Pro Tyr Pro Leu Val Gln Ala Val Lys Arg
405 410 415

Ser Leu Gly Ser Leu
420

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGCGGATCCA TGGGAGCAAC CACCATGGAC CAG 33

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGCGGATCCG CTTCTCTGTA AGTTAATCCT 30

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCGGATCCG CCATCATGGG AGCAACCACC ATG 33

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCGGATCCG CTTCTCTGTA AGTTAATCCT 30

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGCGGATCCG CCATCATGGG AGCAACCACC ATG 33

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
CGCGGATCCT CAAGCGTAGT CTGGGACGTC GTATGGGTAC AGGGAGCCAA GGCTTCT          57
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1432 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 27..1181

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
GGAATTCGGC ACGAGGAGCA ACCACC ATG GAC CAG AAG TCT CTC TGG GCA GGT        53
                              Met Asp Gln Lys Ser Leu Trp Ala Gly
                                1               5

GTA GTG GTC TTG CTG CTT CTC CAG GGA GGA TCT GCC TAC AAA CTG GTT        101
Val Val Val Leu Leu Leu Leu Gln Gly Gly Ser Ala Tyr Lys Leu Val
 10                  15                  20                  25

TGC TAC TTT ACC AAC TGG TCC CAG GAC CGG CAG GAA CCA GGA AAA TTC        149
Cys Tyr Phe Thr Asn Trp Ser Gln Asp Arg Gln Glu Pro Gly Lys Phe
                 30                  35                  40

ACC CCT GAG AAT ATT GAC CCC TTC CTA TGC TCT CAT CTC ATC TAT TCA        197
Thr Pro Glu Asn Ile Asp Pro Phe Leu Cys Ser His Leu Ile Tyr Ser
             45                  50                  55

TTC GCC AGC ATC GAA AAC AAC AAG GTT ATC ATC AAG GAC AAG AGT GAA        245
Phe Ala Ser Ile Glu Asn Asn Lys Val Ile Ile Lys Asp Lys Ser Glu
         60                  65                  70

GTG ATG CTC TAC CAG ACC ATC AAC AGT CTC AAA ACC AAG AAT CCC AAA        293
Val Met Leu Tyr Gln Thr Ile Asn Ser Leu Lys Thr Lys Asn Pro Lys
     75                  80                  85

CTG AAA ATT CTC TTG TCC ATT GGA GGG TAC CTG TTT GGT TCC AAA GGG        341
Leu Lys Ile Leu Leu Ser Ile Gly Gly Tyr Leu Phe Gly Ser Lys Gly
 90                  95                  100                 105

TTC CAC CCT ATG GTG GAT TCT TCT ACA TCA CGC TTG GAA TTC ATT AAC        389
Phe His Pro Met Val Asp Ser Ser Thr Ser Arg Leu Glu Phe Ile Asn
                110                 115                 120
```

```
TCC ATA ATC CTG TTT CTG AGG AAC CAT AAC TTT GAT GGA CTG GAT GTA          437
Ser Ile Ile Leu Phe Leu Arg Asn His Asn Phe Asp Gly Leu Asp Val
            125                 130                 135

AGC TGG ATC TAC CCA GAT CAG AAA GAA AAC ACT CAT TTC ACT GTG CTG          485
Ser Trp Ile Tyr Pro Asp Gln Lys Glu Asn Thr His Phe Thr Val Leu
        140                 145                 150

ATT CAT GAG TTA GCA GAA GCC TTT CAG AAG GAC TTC ACA AAA TCC ACC          533
Ile His Glu Leu Ala Glu Ala Phe Gln Lys Asp Phe Thr Lys Ser Thr
    155                 160                 165

AAG GAA AGG CTT CTC TTG ACT GCG GGC GTA TCT GCA GGG AGG CAA ATG          581
Lys Glu Arg Leu Leu Leu Thr Ala Gly Val Ser Ala Gly Arg Gln Met
170                 175                 180                 185

ATT GAT AAC AGC TAT CAA GTT GAG AAA CTG GCA AAA GAT CTG GAT TTC          629
Ile Asp Asn Ser Tyr Gln Val Glu Lys Leu Ala Lys Asp Leu Asp Phe
                190                 195                 200

ATC AAC CTC CTG TCC TTT GAC TTC CAT GGG TCT TGG GAA AAG CCC CTT          677
Ile Asn Leu Leu Ser Phe Asp Phe His Gly Ser Trp Glu Lys Pro Leu
            205                 210                 215

ATC ACT GGC CAC AAC AGC CCT CTG AGC AAG GGG TGG CAG GAC AGA GGG          725
Ile Thr Gly His Asn Ser Pro Leu Ser Lys Gly Trp Gln Asp Arg Gly
        220                 225                 230

CCA AGC TCC TAC TAC AAT GTG GAA TAT GCT GTG GGG TAC TGG ATA CAT          773
Pro Ser Ser Tyr Tyr Asn Val Glu Tyr Ala Val Gly Tyr Trp Ile His
    235                 240                 245

AAG GGA ATG CCA TCA GAG AAG GTG GTC ATG GGC ATC CCC ACA TAT GGG          821
Lys Gly Met Pro Ser Glu Lys Val Val Met Gly Ile Pro Thr Tyr Gly
250                 255                 260                 265

CAC TCC TTC ACA CTG GCC TCT GCA GAA ACC ACC GTG GGG GCC CCT GCC          869
His Ser Phe Thr Leu Ala Ser Ala Glu Thr Thr Val Gly Ala Pro Ala
                270                 275                 280

TCT GGC CCT GGA GCT GCT GGA CCC ATC ACA GAG TCT TCA GGC TTC CTG          917
Ser Gly Pro Gly Ala Ala Gly Pro Ile Thr Glu Ser Ser Gly Phe Leu
            285                 290                 295

GCC TAT TAT GAG ATC TGC CAG TTC CTG AAA GGA GCC AAG ATC ACG AGG          965
Ala Tyr Tyr Glu Ile Cys Gln Phe Leu Lys Gly Ala Lys Ile Thr Arg
        300                 305                 310

CTC CAG GAT CAG CAG GTT CCC TAC GCA GTC AAG GGG AAC CAG TGG GTG         1013
Leu Gln Asp Gln Gln Val Pro Tyr Ala Val Lys Gly Asn Gln Trp Val
    315                 320                 325

GGC TAT GAT GAT GTG AAG AGT ATG GAG ACC AAG GTT CAG TTC TTA AAG         1061
Gly Tyr Asp Asp Val Lys Ser Met Glu Thr Lys Val Gln Phe Leu Lys
330                 335                 340                 345
```

```
AAT TTA AAC CTG GGA GGA GCC ATG ATC TGG TCT ATT GAC ATG GAT GAC      1109
Asn Leu Asn Leu Gly Gly Ala Met Ile Trp Ser Ile Asp Met Asp Asp
                350                 355                 360

TTC ACT GGC AAA TCC TGC AAC CAG GGC CCT TAC CCT CTT GTC CAA GCA      1157
Phe Thr Gly Lys Ser Cys Asn Gln Gly Pro Tyr Pro Leu Val Gln Ala
            365                 370                 375

GTC AAG AGA AGC CTT GGC TCC CTG TGAAGGATTA ACTTACAGAG AAGCAGGCAA     1211
Val Lys Arg Ser Leu Gly Ser Leu
        380                 385

GATGACCTTG CTGCCTGGGG CCTGCTCTCT CCCAGGAATT CTCATGTGGG ATTCCCCTTG    1271

CCAGGCCGGC CTTTGGATCT CTCTTCCAAG CCTTTCCTGA CTTCCTCTTA GATCATAGAT    1331

TGGACCTGGT TTTGTTTTCC TGCAGCTGAT GACTTGTTGC CCTGAAGTAC AATAAAAAAA    1391

ATTCATTTTG CTCCAAAAAA AAAAAAAAAA AAAAACTCGA G                        1432
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 385 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Met Asp Gln Lys Ser Leu Trp Ala Gly Val Val Leu Leu Leu Leu
 1               5                  10                  15

Gln Gly Gly Ser Ala Tyr Lys Leu Val Cys Tyr Phe Thr Asn Trp Ser
            20                  25                  30

Gln Asp Arg Gln Glu Pro Gly Lys Phe Thr Pro Glu Asn Ile Asp Pro
        35                  40                  45

Phe Leu Cys Ser His Leu Ile Tyr Ser Phe Ala Ser Ile Glu Asn Asn
    50                  55                  60

Lys Val Ile Ile Lys Asp Lys Ser Glu Val Met Leu Tyr Gln Thr Ile
65                  70                  75                  80

Asn Ser Leu Lys Thr Lys Asn Pro Lys Leu Lys Ile Leu Leu Ser Ile
                85                  90                  95

Gly Gly Tyr Leu Phe Gly Ser Lys Gly Phe His Pro Met Val Asp Ser
            100                 105                 110

Ser Thr Ser Arg Leu Glu Phe Ile Asn Ser Ile Ile Leu Phe Leu Arg
        115                 120                 125
```

```
Asn His Asn Phe Asp Gly Leu Asp Val Ser Trp Ile Tyr Pro Asp Gln
    130                 135                 140

Lys Glu Asn Thr His Phe Thr Val Leu Ile His Glu Leu Ala Glu Ala
145                 150                 155                 160

Phe Gln Lys Asp Phe Thr Lys Ser Thr Lys Glu Arg Leu Leu Leu Thr
                165                 170                 175

Ala Gly Val Ser Ala Gly Arg Gln Met Ile Asp Asn Ser Tyr Gln Val
            180                 185                 190

Glu Lys Leu Ala Lys Asp Leu Asp Phe Ile Asn Leu Leu Ser Phe Asp
        195                 200                 205

Phe His Gly Ser Trp Glu Lys Pro Leu Ile Thr Gly His Asn Ser Pro
    210                 215                 220

Leu Ser Lys Gly Trp Gln Asp Arg Gly Pro Ser Ser Tyr Tyr Asn Val
225                 230                 235                 240

Glu Tyr Ala Val Gly Tyr Trp Ile His Lys Gly Met Pro Ser Glu Lys
                245                 250                 255

Val Val Met Gly Ile Pro Thr Tyr Gly His Ser Phe Thr Leu Ala Ser
            260                 265                 270

Ala Glu Thr Thr Val Gly Ala Pro Ala Ser Gly Pro Gly Ala Ala Gly
        275                 280                 285

Pro Ile Thr Glu Ser Ser Gly Phe Leu Ala Tyr Tyr Glu Ile Cys Gln
    290                 295                 300

Phe Leu Lys Gly Ala Lys Ile Thr Arg Leu Gln Asp Gln Gln Val Pro
305                 310                 315                 320

Tyr Ala Val Lys Gly Asn Gln Trp Val Gly Tyr Asp Asp Val Lys Ser
                325                 330                 335

Met Glu Thr Lys Val Gln Phe Leu Lys Asn Leu Asn Leu Gly Gly Ala
            340                 345                 350

Met Ile Trp Ser Ile Asp Met Asp Asp Phe Thr Gly Lys Ser Cys Asn
        355                 360                 365

Gln Gly Pro Tyr Pro Leu Val Gln Ala Val Lys Arg Ser Leu Gly Ser
    370                 375                 380

Leu
385
```

70.1

Applicant's or agent's file reference number 1488.04[illegible]00	International application No TB[illegible]

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 5, line 16.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution

American Type Culture Collection

Address of depositary institution *(including postal code and country)*

12301 Parklawn Drive
Rockville, Maryland 20852
United State of America

Date of deposit	Accession Number
May 3, 1996	ATCC Designation 97529

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

DNA plasmid 1290853 (Docket PF286)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

Form PCT/RO/134 (July 1992)

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a nucleotide sequence encoding the chitinase alpha polypeptide having the complete amino acid sequence in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4];

(b) a nucleotide sequence encoding the mature chitinase alpha polypeptide having the amino acid sequence at positions 27 - 421 in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4];

(c) a nucleotide sequence encoding the chitinase alpha polypeptide having the complete amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97529;

(d) a nucleotide sequence encoding the mature chitinase alpha polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97529;

(e) a nucleotide sequence encoding the chitinase alpha-2 polypeptide having the complete amino acid sequence shown in Figure 5 [SEQ ID NO:12];

(f) a nucleotide sequence encoding the mature chitinase alpha-2 polypeptide having the amino acid sequence at positions 21-385 in Figure 5 [SEQ ID NO:12]; and

(g) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e), or (f).

2. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence in Figure 1 [SEQ ID NO:1], Figure 2 [SEQ ID NO:3], or Figure 5 [SEQ ID NO:11].

3. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence in Figure 1 [SEQ ID NO:1], Figure 2 [SEQ ID NO:3], or Figure 5 [SEQ ID NO:11] encoding the polypeptide having the complete amino acid sequence in Figure 1 [SEQ ID NO:2], Figure 2 [SEQ ID NO:4], or Figure 5 [SEQ ID NO:12], respectively.

4. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence in Figure 1 [SEQ ID NO:1], Figure 2 [SEQ ID NO:3], or Figure 5 [SEQ ID NO:11] encoding the mature chitinase alpha polypeptide having the amino acid sequence at positions 27 - 421 in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4] or encoding the mature chitinase alpha-2 polypeptide having the amino acid sequence at positions 21-385 in Figure 5 [SEQ ID NO:12].

5. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence of the cDNA clone contained in ATCC Deposit No. 97529.

6. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence encoding the chitinase alpha polypeptide having the complete amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97529.

7. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence encoding the mature chitinase alpha polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97529.

8. An isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide having a nucleotide sequence identical to a nucleotide sequence in (a), (b), (c), (d),

(e), (f), or (g) of claim 1 wherein said polynucleotide which hybrizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues.

9. An isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of a chitinase alpha or chitinase alpha-2 polypeptide having an amino acid sequence in (a), (b), (c), (d), (e), or (f) of claim 1.

10. The isolated nucleic acid molecule of claim 9, which encodes an epitope-bearing portion of a chitinase alpha or chitinase alpha-2 polypeptide selected from the group consisting of: a polypeptide comprising amino acid residues from about 38 to about 51 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 66 to about 85 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 117 to about 127 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 145 to about 153 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 165 to about 172 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 177 to about 185 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 245 to about 254 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 256 to about 274 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 286 to about 293 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 366 to about 376 in Figure 1 (SEQ ID NO:2); and a polypeptide comprising amino acid residues from about 394 to about 407 in Figure 1 (SEQ ID NO:2).

11. A method for making a recombinant vector comprising inserting an isolated nucleic acid molecule of claim 1 into a vector.

12. A recombinant vector produced by the method of claim 11.

13. A method of making a recombinant host cell comprising introducing the recombinant vector of claim 12 into a host cell.

14. A recombinant host cell produced by the method of claim 13.

15. A recombinant method for producing a chitinase alpha or chitinase alpha-2 polypeptide, comprising culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed and recovering said polypeptide.

16. An isolated chitinase alpha or chitinase alpha-2 polypeptide having an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) the amino acid sequence of the chitinase alpha polypeptide having the complete amino acid sequence in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4];

(b) the amino acid sequence of the mature chitinase alpha polypeptide having the amino acid sequence at positions 27 - 421 in Figure 1 [SEQ ID NO:2] or Figure 2 [SEQ ID NO:4];

(c) the amino acid sequence of the chitinase alpha polypeptide having the complete amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97529; and

(d) the amino acid sequence of the mature chitinase alpha polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97529;

(e) the amino acid sequence of the chitinase alpha-2 polypeptide having the complete 385 amino acid sequence including the leader sequence shown in Figure 5 [SEQ ID NO:12];

(f) the amino acid sequence of the mature chitinase alpha-2 polypeptide (without the leader) having the amino acid sequence at positions 21-385 in Figure 5 [SEQ ID NO:12]. and

(e) the amino acid sequence of an epitope-bearing portion of any one of the polypeptides of (a), (b), (c), (d), (e), or (f).

17. An isolated polypeptide comprising an epitope-bearing portion of chitinase alpha, wherein said portion is selected from the group consisting of: a polypeptide comprising amino acid residues from about 38 to about 51 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 66 to about 85 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 117 to about 127 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 145 to about 153 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 165 to about 172 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 177 to about 185 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 245 to about 254 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 256 to about 274 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 286 to about 293 in Figure 1 (SEQ ID NO:2); a polypeptide comprising amino acid residues from about 366 to about 376 in Figure 1 (SEQ ID NO:2); and a polypeptide comprising amino acid residues from about 394 to about 407 in Figure 1 (SEQ ID NO:2).

18. An isolated antibody that binds specifically to a chitinase alpha or chitinase alpha-2 polypeptide of claim 16.

19. A method for treating a patient having need of anti-fungal therapy comprising administering to said patient a therapeutically effective amount of the polypeptide of claim 16.

20. A method for treating a patient having a tissue remodeling disorder, comprising administering to said patient a therapeutically effective amount of an antagonist of the polypeptide of claim 16.

```
        10                  30                  50
gaattcggcacgagggctgtcgaaacctcagtggataaaagacctagagaatgtgtatcc
        70                  90                  110
cagaagaagctggccaaggatATGGGAGCAACCACCATGGACCAGAAGTCTCTCTGGGCA
                            M  G  A  T  T  M  D  Q  K  S  L  W  A
       130                 150                 170
GGTGTAGTGGTCTTGCTGCTTCTCCAGGGAGAGATGGGGTTTTGCTATGTTGCCAGAGCT
G  V  V  V  L  L  L  L  Q  G  E  M  G  F  C  Y  V  A  R  A
       190                 210                 230
GGTCTTGAACTCCTGGGCTCAAGAAGTCCTCCTGCCTCAGCCTCCCAAAGTGCTGGGATA
G  L  E  L  L  G  S  R  S  P  P  A  S  A  S  Q  S  A  G  I
       250                 270                 290
ACAGGATCTGCCTACAAACTGGTTTGCTACTTTACCAACTGGTCCCAGGACCGGCAGGAA
T  G  S  A  Y  K  L  V  C  Y  F  T  N  W  S  Q  D  R  Q  E
       310                 330                 350
CCAGGAAAATTCACCCCTGAGAATATTGACCCCTTCCTATGCTCTCATCTCATCTATTCA
P  G  K  F  T  P  E  N  I  D  P  F  L  C  S  H  L  I  Y  S
       370                 390                 410
TTCGCCAGCATCGAAAACAACAAGGTTATCATCAAGGACAAGAGTGAAGTGATGCTCTAC
F  A  S  I  E  N  N  K  V  I  I  K  D  K  S  E  V  M  L  Y
       430                 450                 470
CAGACCATCAACAGTCTCAAAACCAAGAATCCCAAACTGAAAATTCTCTTGTCCATTGGA
Q  T  I  N  S  L  K  T  K  N  P  K  L  K  I  L  L  S  I  G
       490                 510                 530
GGGTACCTGTTTGGTTCCAAAGGGTTCCACCCTATGGTGGATTCTTCTACATCACGCTTG
G  Y  L  F  G  S  K  G  F  H  P  M  V  D  S  S  T  S  R  L
       550                 570                 590
GAATTCATTAACTCCATAATCCTGTTTCTGAGGAACCATAACTTTGATGGACTGGATGTA
E  F  I  N  S  I  I  L  F  L  R  N  H  N  F  D  G  L  D  V
       610                 630                 650
AGCTGGATCTACCCAGATCAGAAAGAAAACACTCATTTCACTGTGCTGATTCATGAGTTA
S  W  I  Y  P  D  Q  K  E  N  T  H  F  T  V  L  I  H  E  L
       670                 690                 710
GCAGAAGCCTTTCAGAAGGACTTCACAAAATCCACCAAGGAAAGGCTTCTCTTGACTGCG
A  E  A  F  Q  K  D  F  T  K  S  T  K  E  R  L  L  L  T  A
       730                 750                 770
GGCGTATCTGCAGGGAGGCAAATGATTGATAACAGCTATCAAGTTGAGAAACTGGCAAAA
G  V  S  A  G  R  Q  M  I  D  N  S  Y  Q  V  E  K  L  A  K
       790                 810                 830
GATCTGGATTTCATCAACCTCCTGTCCTTTGACTTCCATGGGTCTTGGGAAAAGCCCCTT
D  L  D  F  I  N  L  L  S  F  D  F  H  G  S  W  E  K  P  L
```

FIG.1A

```
       850                 870                 890
ATCACTGGCCACAACAGCCCTCTGAGCAAGGGGTGGCAGGACAGAGGGCCAAGCTCCTAC
I  T  G  H  N  S  P  L  S  K  G  W  Q  D  R  G  P  S  S  Y
       910                 930                 950
TACAATGTGGAATATGCTGTGGGGTACTGGATACATAAGGGAATGCCATCAGAGAAGGTG
Y  N  V  E  Y  A  V  G  Y  W  I  H  K  G  M  P  S  E  K  V
       970                 990                1010
GTCATGGGCATCCCCACATATGGGCACTCCTTCACACTGGCCTCTGCAGAAACCACCGTG
V  M  G  I  P  T  Y  G  H  S  F  T  L  A  S  A  E  T  T  V
      1030                1050                1070
GGGGCCCCTGCCTCTGGCCCTGGAGCTGCTGGACCCATCACAGAGTCTTCAGGCTTCCTG
G  A  P  A  S  G  P  G  A  A  G  P  I  T  E  S  S  G  F  L
      1090                1110                1130
GCCTATTATGAGATCTGCCAGTTCCTGAAAGGAGCCAAGATCACGAGGCTCCAGGATCAG
A  Y  Y  E  I  C  Q  F  L  K  G  A  K  I  T  R  L  Q  D  Q
      1150                1170                1190
CAGGTTCCCTACGCAGTCAAGGGGAACCAGTGGGTGGGCTATGATGATGTGAAGAGTATG
Q  V  P  Y  A  V  K  G  N  Q  W  V  G  Y  D  D  V  K  S  M
      1210                1230                1250
GAGACCAAGGTTCAGTTCTTAAAGAATTTAAACCTGGGAGGAGCCATGATCTGGTCTATT
E  T  K  V  Q  F  L  K  N  L  N  L  G  G  A  M  I  W  S  I
      1270                1290                1310
GACATGGATGACTTCACTGGCAAATCCTGCAACCAGGGCCCTTACCCTCTTGTCCAAGCA
D  M  D  D  F  T  G  K  S  C  N  Q  G  P  Y  P  L  V  Q  A
      1330                1350                1370
GTCAAGAGAAGCCTTGGCTCCCTGTGAaggattaacttacagagaagcaggcaagatgac
V  K  R  S  L  G  S  L  *
      1390                1410                1430
cttgctgcctggggcctgctctctcccaggaattctcatgtgggattccccttgccaggc
      1450                1470                1490
cggcctttggatctctcttccaagcctttcctgacttcctcttagatcatagattggacc
      1510                1530                1550
tggttttgttttcctgcagctgatgacttgttgccctgaagtacaataaaaaaaattcat
      1570                1590
tttgctccagtaaaaaaaaaaaaaaaaactcgag
```

FIG.1B

```
        10                  30                  50
gaattcggcacgagggctgtcgaaacctcagtggataaaagacctagagaatgtgtatcc
        70                  90                  110
cagaaggaagctggccaaggatATGGGAGCAACCACCATGGACCAGAAGTCTCTCTGGGCA
                            M  G  A  T  T  M  D  Q  K  S  L  W  A
       130                 150                 170
GGTGTAGTGGTCTTGCTGCTTCTCCAGGGAGAGATGGGGTTTTGCTATGTTGCCAGAGCT
G  V  V  V  L  L  L  L  Q  G  E  M  G  F  C  Y  V  A  R  A
       190                 210                 230
GGTCTTGAACTCCTGGGCTCAAGAAGTCCTCCTGCCTCAGCCTCCCAAAGTGCTGGGATA
G  L  E  L  L  G  S  R  S  P  P  A  S  A  S  Q  S  A  G  I
       250                 270                 290
ACAGGATCTGCCTACAAACTGGTTTGCTACTTTACCAACTGGTCCCAGGACCGGCAGGAA
T  G  S  A  Y  K  L  V  C  Y  F  T  N  W  S  Q  D  R  Q  E
       310                 330                 350
CCAGGAAAATTCACCCCTGAGAATATTGACCCCTTCCTATGCTCTCATCTCATCTATTCA
P  G  K  F  T  P  E  N  I  D  P  F  L  C  S  H  L  I  Y  S
       370                 390                 410
TTCGCCAGCATCGAAAACAACAAGGTTATCATCAAGGACAAGAGTGAAGTGATGCTCTAC
F  A  S  I  E  N  N  K  V  I  I  K  D  K  S  E  V  M  L  Y
       430                 450                 470
CAGACCATCAACAGTCTCAAAACCAAGAATCCCAAACTGAAAATTCTCTTGTCCATTGGA
Q  T  I  N  S  L  K  T  K  N  P  K  L  K  I  L  L  S  I  G
       490                 510                 530
GGGTACCTGTTTGGTTCCAAAGGGTTCCACCCTATGGTGGATTCTTCTACATCACGCTTG
G  Y  L  F  G  S  K  G  F  H  P  M  V  D  S  S  T  S  R  L
       550                 570                 590
GAATTCATTAACTCCATAATCCTGTTTCTGAGGAACCATAACTTTGATGGACTGGATGTA
E  F  I  N  S  I  I  L  F  L  R  N  H  N  F  D  G  L  D  V
       610                 630                 650
AGCTGGGAATACCCAGATCAGAAAGAAAACACTCATTTCACTGTGCTGATTCATGAGTTA
S  W  E  Y  P  D  Q  K  E  N  T  H  F  T  V  L  I  H  E  L
       670                 690                 710
GCAGAAGCCTTTCAGAAGGACTTCACAAAATCCACCAAGGAAAGGCTTCTCTTGACTGCG
A  E  A  F  Q  K  D  F  T  K  S  T  K  E  R  L  L  L  T  A
       730                 750                 770
GGCGTATCTGCAGGGAGGCAAATGATTGATAACAGCTATCAAGTTGAGAAACTGGCAAAA
G  V  S  A  G  R  Q  M  I  D  N  S  Y  Q  V  E  K  L  A  K
       790                 810                 830
GATCTGGATTTCATCAACCTCCTGTCCTTTGACTTCCATGGGTCTTGGGAAAAGCCCCTT
D  L  D  F  I  N  L  L  S  F  D  F  H  G  S  W  E  K  P  L
```

FIG.2A

```
       850                 870                 890
ATCACTGGCCACAACAGCCCTCTGAGCAAGGGGTGGCAGGACAGAGGGCCAAGCTCCTAC
I  T  G  H  N  S  P  L  S  K  G  W  Q  D  R  G  P  S  S  Y
       910                 930                 950
TACAATGTGGAATATGCTGTGGGGTACTGGATACATAAGGGAATGCCATCAGAGAAGGTG
Y  N  V  E  Y  A  V  G  Y  W  I  H  K  G  M  P  S  E  K  V
       970                 990                1010
GTCATGGGCATCCCCACATATGGGCACTCCTTCACACTGGCCTCTGCAGAAACCACCGTG
V  M  G  I  P  T  Y  G  H  S  F  T  L  A  S  A  E  T  T  V
      1030                1050                1070
GGGGCCCCTGCCTCTGGCCCTGGAGCTGCTGGACCCATCACAGAGTCTTCAGGCTTCCTG
G  A  P  A  S  G  P  G  A  A  G  P  I  T  E  S  S  G  F  L
      1090                1110                1130
GCCTATTATGAGATCTGCCAGTTCCTGAAAGGAGCCAAGATCACGAGGCTCCAGGATCAG
A  Y  Y  E  I  C  Q  F  L  K  G  A  K  I  T  R  L  Q  D  Q
      1150                1170                1190
CAGGTTCCCTACGCAGTCAAGGGGAACCAGTGGGTGGGCTATGATGATGTGAAGAGTATG
Q  V  P  Y  A  V  K  G  N  Q  W  V  G  Y  D  D  V  K  S  M
      1210                1230                1250
GAGACCAAGGTTCAGTTCTTAAAGAATTTAAACCTGGGAGGAGCCATGATCTGGTCTATT
E  T  K  V  Q  F  L  K  N  L  N  L  G  G  A  M  I  W  S  I
      1270                1290                1310
GACATGGATGACTTCACTGGCAAATCCTGCAACCAGGGCCCTTACCCTCTTGTCCAAGCA
D  M  D  D  F  T  G  K  S  C  N  Q  G  P  Y  P  L  V  Q  A
      1330                1350                1370
GTCAAGAGAAGCCTTGGCTCCCTGTGAaggattaacttacagagaagcaggcaagatgac
V  K  R  S  L  G  S  L  *
      1390                1410                1430
cttgctgcctggggcctgctctctcccaggaattctcatgtgggattccccttgccaggc
      1450                1470                1490
cggcctttggatctctcttccaagcctttcctgacttcctcttagatcatagattggacc
      1510                1530                1550
tggttttgttttcctgcagctgatgacttgttgccctgaagtacaataaaaaaaattcat
      1570                1590
tttgctccagtaaaaaaaaaaaaaaaaactcgag
```

FIG.2B

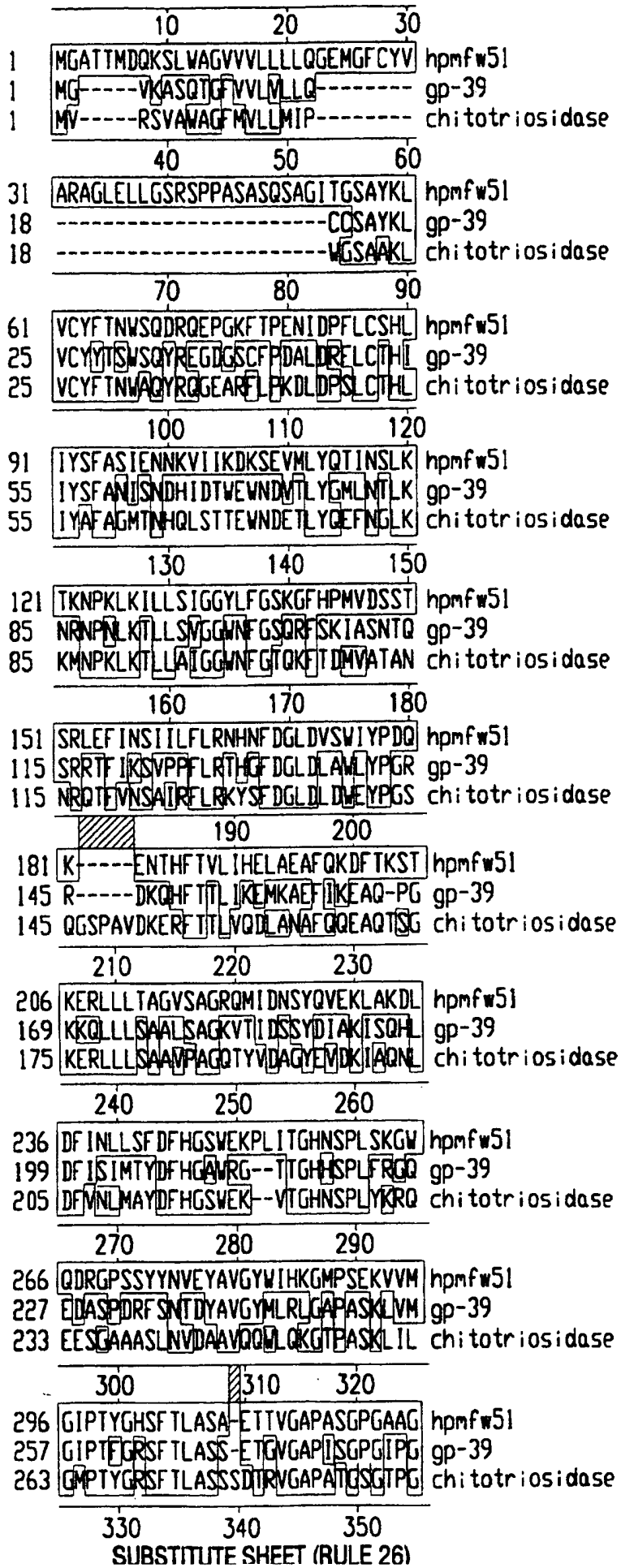
10 20 30
1 MGATTMDQKSLWAGVVVLLLLQGEMGFCYV hpmfw51
1 MG-----VKASQTGFVVLVLLQ-------- gp-39
1 MV-----RSVAWAGFMVLLMIP-------- chitotriosidase
40 50 60
31 ARAGLELLGSRSPPASASQSAGITGSAYKL hpmfw51
18 ----------CQSAYKL gp-39
18 ----------WGSAAKL chitotriosidase
70 80 90
61 VCYFTNWSQDRQEPGKFTPENIDPFLCSHL hpmfw51
25 VCYYTSWSQYREGDGSCFPDALDRFLCTHL gp-39
25 VCYFTNWAQYRQGEARFLPKDLDPSLCTHL chitotriosidase
100 110 120
91 IYSFASIENNKVIIKDKSEVMLYQTINSLK hpmfw51
55 IYSFANISNDHIDTWEWNDVTLYGMLNTLK gp-39
55 IYAFAGMTNHQLSTTEWNDETLYQEFNGLK chitotriosidase
130 140 150
121 TKNPKLKILLSIGGYLFGSKGFHPMVDSST hpmfw51
85 NRNPNLKTLLSVGGWNFGSQRFSKIASNTQ gp-39
85 KMNPKLKTLLAIGGWNFGTQKFTDMVATAN chitotriosidase
160 170 180
151 SRLEFINSIILFLRNHNFDGLDVSWIYPDQ hpmfw51
115 SRRTFIKSVPPFLRTHGFDGLDLAWLYPGR gp-39
115 NRQTFVNSAIRFLRKYSFDGLDLDWEYPGS chitotriosidase
190 200
181 K-----ENTHFTVLIHELAEAFQKDFTKST hpmfw51
145 R-----DKQHFTTLIKEMKAEFIKEAQ-PG gp-39
145 QGSPAVDKERFTTLVQDLANAFQQEAQTSG chitotriosidase
210 220 230
206 KERLLLTAGVSAGRQMIDNSYQVEKLAKDL hpmfw51
169 KKQLLLSAALSAGKVTIDSSYDIAKISQHL gp-39
175 KERLLLSAAMPAGQTYVDAGYEVDKIAQNL chitotriosidase
240 250 260
236 DFINLLSFDFHGSWEKPLITGHNSPLSKGW hpmfw51
199 DFISIMTYDFHGAWRG--TTGHHSPLFRGQ gp-39
205 DFVNLMAYDFHGSWEK--VTGHNSPLYKRQ chitotriosidase
270 280 290
266 QDRGPSSYYNVEYAVGYWIHKGMPSEKVVM hpmfw51
227 EDASPDRFSNTDYAVGYMLRLGAPASKLVM gp-39
233 EESGAAASLNVDAAVQQWLQKGTPASKLIL chitotriosidase
300 310 320
296 GIPTYGHSFTLASA-ETTVGAPASGPGAAG hpmfw51
257 GIPTFGRSFTLASS-ETGVGAPISGPGIPG gp-39
263 GMPTYGRSFTLASSSDTRVGAPATGSGTPG chitotriosidase
330 340 350

FIG.3

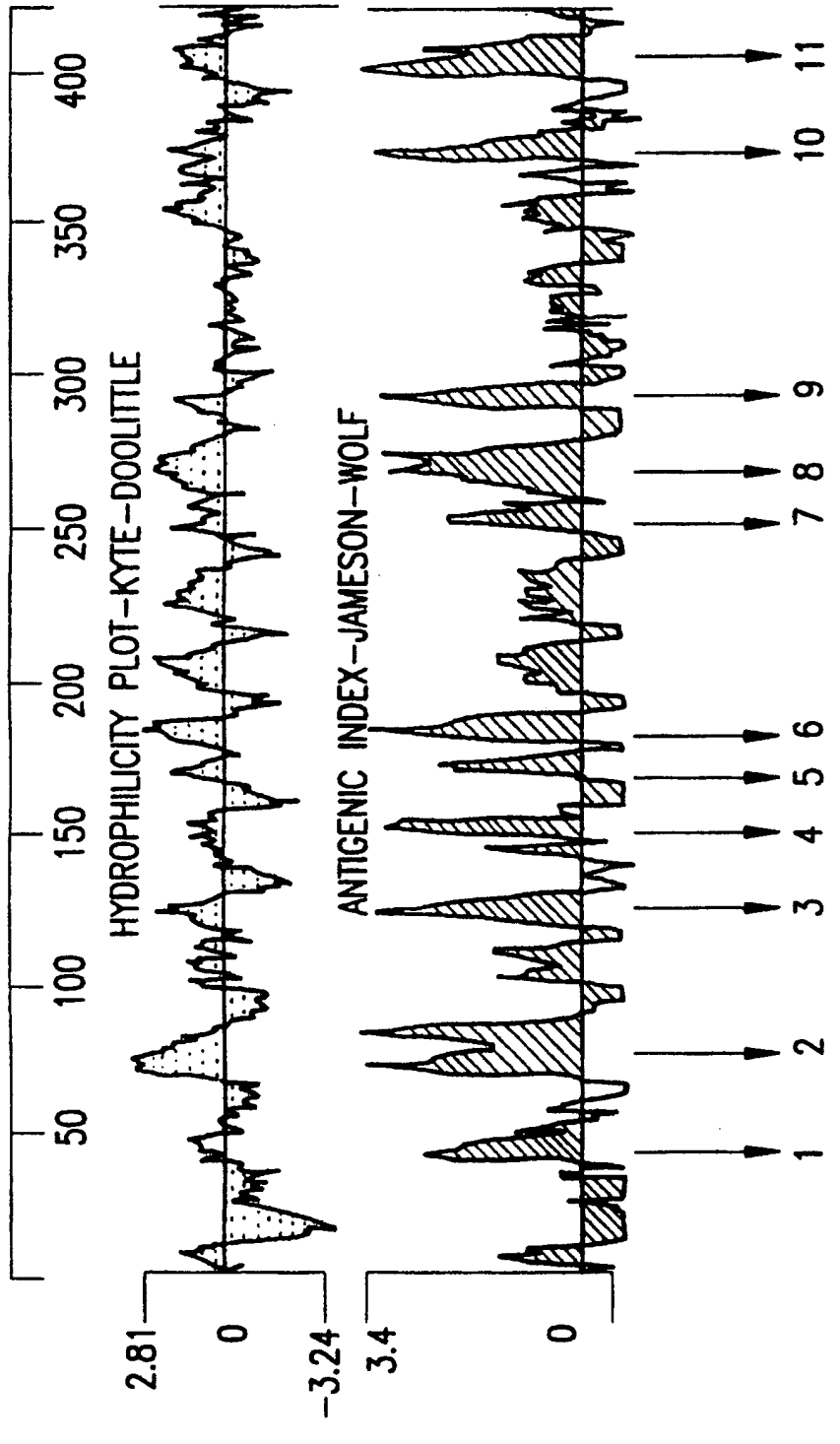
HYDROPHILICITY PLOT-KYTE-DOOLITTLE
50
100
150
200
250
300
350
400
2.81
0
-3.24
ANTIGENIC INDEX-JAMESON-WOLF
3.4
0
1
2
3
4
5
6
7
8
9
10
11
ANTIGENIC PEPTIDES:
1. Leu38-Ala51
2. Asn66-Phe85
3. Asn117-Lys127
4. Met145-Leu153
5. Asn165-Asp172
6. Tyr177-His185
7. Phe245-Ile254
8. Gly256-Tyr274
9. Lys286-Tyr293
10. Gly366-Lys376
11.Asp394-Pro407

FIG.4

```
        10                  30                  50
GGAATTCGGCACGAGGAGCAACCACCATGGACCAGAAGTCTCTCTGGGCAGGTGTAGTGG
                            M  D  Q  K  S  L  W  A  G  V  V  V
        70                  90                  110
TCTTGCTGCTTCTCCAGGGAGGATCTGCCTACAAACTGGTTTGCTACTTTACCAACTGGT
  L  L  L  L  Q  G  G  S  A  Y  K  L  V  C  Y  F  T  N  W  S
       130                 150                 170
CCCAGGACCGGCAGGAACCAGGAAAATTCACCCCTGAGAATATTGACCCCTTCCTATGCT
  Q  D  R  Q  E  P  G  K  F  T  P  E  N  I  D  P  F  L  C  S
       190                 210                 230
CTCATCTCATCTATTCATTCGCCAGCATCGAAAACAACAAGGTTATCATCAAGGACAAGA
  H  L  I  Y  S  F  A  S  I  E  N  N  K  V  I  I  K  D  K  S
       250                 270                 290
GTGAAGTGATGCTCTACCAGACCATCAACAGTCTCAAAACCAAGAATCCCAAACTGAAAA
  E  V  M  L  Y  Q  T  I  N  S  L  K  T  K  N  P  K  L  K  I
       310                 330                 350
TTCTCTTGTCCATTGGAGGGTACCTGTTTGGTTCCAAAGGGTTCCACCCTATGGTGGATT
  L  L  S  I  G  G  Y  L  F  G  S  K  G  F  H  P  M  V  D  S
       370                 390                 410
CTTCTACATCACGCTTGGAATTCATTAACTCCATAATCCTGTTTCTGAGGAACCATAACT
  S  T  S  R  L  E  F  I  N  S  I  I  L  F  L  R  N  H  N  F
       430                 450                 470
TTGATGGACTGGATGTAAGCTGGATCTACCCAGATCAGAAAGAAAACACTCATTTCACTG
  D  G  L  D  V  S  W  I  Y  P  D  Q  K  E  N  T  H  F  T  V
       490                 510                 530
TGCTGATTCATGAGTTAGCAGAAGCCTTTCAGAAGGACTTCACAAAATCCACCAAGGAAA
  L  I  H  E  L  A  E  A  F  Q  K  D  F  T  K  S  T  K  E  R
       550                 570                 590
GGCTTCTCTTGACTGCGGGCGTATCTGCAGGGAGGCAAATGATTGATAACAGCTATCAAG
  L  L  L  T  A  G  V  S  A  G  R  Q  M  I  D  N  S  Y  Q  V
       610                 630                 650
TTGAGAAACTGGCAAAAGATCTGGATTTCATCAACCTCCTGTCCTTTGACTTCCATGGGT
  E  K  L  A  K  D  L  D  F  I  N  L  L  S  F  D  F  H  G  S
       670                 690                 710
CTTGGGAAAAGCCCCTTATCACTGGCCACAACAGCCCTCTGAGCAAGGGGTGGCAGGACA
  W  E  K  P  L  I  T  G  H  N  S  P  L  S  K  G  W  Q  D  R
       730                 750                 770
GAGGGCCAAGCTCCTACTACAATGTGGAATATGCTGTGGGGTACTGGATACATAAGGGAA
  G  P  S  S  Y  Y  N  V  E  Y  A  V  G  Y  W  I  H  K  G  M
       790                 810                 830
TGCCATCAGAGAAGGTGGTCATGGGCATCCCCACATATGGGCACTCCTTCACACTGGCCT
  P  S  E  K  V  V  M  G  I  P  T  Y  G  H  S  F  T  L  A  S
```

FIG.5A

```
      850                 870                 890
CTGCAGAAACCACCGTGGGGGCCCCTGCCTCTGGCCCTGGAGCTGCTGGACCCATCACAG
  A  E  T  T  V  G  A  P  A  S  G  P  G  A  A  G  P  I  T  E
      910                 930                 950
AGTCTTCAGGCTTCCTGGCCTATTATGAGATCTGCCAGTTCCTGAAAGGAGCCAAGATCA
  S  S  G  F  L  A  Y  Y  E  I  C  Q  F  L  K  G  A  K  I  T
      970                 990                1010
CGAGGCTCCAGGATCAGCAGGTTCCCTACGCAGTCAAGGGGAACCAGTGGGTGGGCTATG
  R  L  Q  D  Q  Q  V  P  Y  A  V  K  G  N  Q  W  V  G  Y  D
     1030                1050                1070
ATGATGTGAAGAGTATGGAGACCAAGGTTCAGTTCTTAAAGAATTTAAACCTGGGAGGAG
  D  V  K  S  M  E  T  K  V  Q  F  L  K  N  L  N  L  G  G  A
     1090                1110                1130
CCATGATCTGGTCTATTGACATGGATGACTTCACTGGCAAATCCTGCAACCAGGGCCCTT
  M  I  W  S  I  D  M  D  D  F  T  G  K  S  C  N  Q  G  P  Y
     1150                1170                1190
ACCCTCTTGTCCAAGCAGTCAAGAGAAGCCTTGGCTCCCTGTGAAGGATTAACTTACAGA
  P  L  V  Q  A  V  K  R  S  L  G  S  L  *
     1210                1230                1250
GAAGCAGGCAAGATGACCTTGCTGCCTGGGGCCTGCTCTCTCCCAGGAATTCTCATGTGG

     1270                1290                1310
GATTCCCCTTGCCAGGCCGGCCTTTGGATCTCTCTTCCAAGCCTTTCCTGACTTCCTCTT

     1330                1350                1370
AGATCATAGATTGGACCTGGTTTTGTTTTCCTGCAGCTGATGACTTGTTGCCCTGAAGTA

     1390                1410                1430
CAATAAAAAAAATTCATTTTGCTCCAAAAAAAAAAAAAAAAAAAACTCGAG
```

FIG.5B

```
  1 ....GGAATTCGGCACGAGGAGCAACCACCATGGACCAGAAGTCTCTCTG 46
        |    ||    ||||||||||||||||||||||||||||||||||
251 aagctggccaaggatATGGGAGCAACCACCATGGACCAGAAGTCTCTCTG 300

 47 GGCAGGTGTAGTGGTCTTGCTGCTTCTCCAGGG..................  79
    |||||||||||||||||||||||||||||||||
301 GGCAGGTGTAGTGGTCTTGCTGCTTCTCCAGGGAGAGATGGGGTTTTGCT 350
    ..................................................

351 ATGTTGCCAGAGCTGGTCTTGAACTCCTGGGCTCAAGAAGTCCTCCTGCC 400

 80 ........................AGGATCTGCCTACAAACTGGTTTG 103
                            |||||||||||||||||||||||
401 TCAGCCTCCCAAAGTGCTGGGATAACAGGATCTGCCTACAAACTGGTTTG 450

104 CTACTTTACCAACTGGTCCCAGGACCGGCAGGAACCAGGAAAATTCACCC 153
    |||||||||||||||||||||||||||||||||||||||||||||||||
451 CTACTTTACCAACTGGTCCCAGGACCGGCAGGAACCAGGAAAATTCACCC 500

154 CTGAGAATATTGACCCCTTCCTATGCTCTCATCTCATCTATTCATTCGCC 203
    |||||||||||||||||||||||||||||||||||||||||||||||||
501 CTGAGAATATTGACCCCTTCCTATGCTCTCATCTCATCTATTCATTCGCC 550

204 AGCATCGAAAACAACAAGGTTATCATCAAGGACAAGAGTGAAGTGATGCT 253
    |||||||||||||||||||||||||||||||||||||||||||||||||
551 AGCATCGAAAACAACAAGGTTATCATCAAGGACAAGAGTGAAGTGATGCT 600

254 CTACCAGACCATCAACAGTCTCAAAACCAAGAATCCCAAACTGAAAATTC 303
    |||||||||||||||||||||||||||||||||||||||||||||||||
601 CTACCAGACCATCAACAGTCTCAAAACCAAGAATCCCAAACTGAAAATTC 650

304 TCTTGTCCATTGGAGGGTACCTGTTTGGTTCCAAAGGGTTCCACCCTATG 353
    |||||||||||||||||||||||||||||||||||||||||||||||||
651 TCTTGTCCATTGGAGGGTACCTGTTTGGTTCCAAAGGGTTCCACCCTATG 700

354 GTGGATTCTTCTACATCACGCTTGGAATTCATTAACTCCATAATCCTGTT 403
    |||||||||||||||||||||||||||||||||||||||||||||||||
701 GTGGATTCTTCTACATCACGCTTGGAATTCATTAACTCCATAATCCTGTT 750

404 TCTGAGGAACCATAACTTTGATGGACTGGATGTAAGCTGGATCTACCCAG 453
    |||||||||||||||||||||||||||||||||||||||||||||||||
751 TCTGAGGAACCATAACTTTGATGGACTGGATGTAAGCTGGATCTACCCAG 800
```

FIG.6A

```
454 ATCAGAAAGAAAACACTCATTTCACTGTGCTGATTCATGAGTTAGCAGAA 503
    ||||||||||||||||||||||||||||||||||||||||||||||||||
801 ATCAGAAAGAAAACACTCATTTCACTGTGCTGATTCATGAGTTAGCAGAA 850

504 GCCTTTCAGAAGGACTTCACAAAATCCACCAAGGAAAGGCTTCTCTTGAC 553
    ||||||||||||||||||||||||||||||||||||||||||||||||||
851 GCCTTTCAGAAGGACTTCACAAAATCCACCAAGGAAAGGCTTCTCTTGAC 900

554 TGCGGGCGTATCTGCAGGGAGGCAAATGATTGATAACAGCTATCAAGTTG 603
    ||||||||||||||||||||||||||||||||||||||||||||||||||
901 TGCGGGCGTATCTGCAGGGAGGCAAATGATTGATAACAGCTATCAAGTTG 950

604 AGAAACTGGCAAAAGATCTGGATTTCATCAACCTCCTGTCCTTTGACTTC 653
    ||||||||||||||||||||||||||||||||||||||||||||||||||
951 AGAAACTGGCAAAAGATCTGGATTTCATCAACCTCCTGTCCTTTGACTTC 1000

654 CATGGGTCTTGGGAAAAGCCCCTTATCACTGGCCACAACAGCCCTCTGAG 703
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1001 CATGGGTCTTGGGAAAAGCCCCTTATCACTGGCCACAACAGCCCTCTGAG 1050

704 CAAGGGGTGGCAGGACAGAGGGCCAAGCTCCTACTACAATGTGGAATATG 753
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1051 CAAGGGGTGGCAGGACAGAGGGCCAAGCTCCTACTACAATGTGGAATATG 1100

754 CTGTGGGGTACTGGATACATAAGGGAATGCCATCAGAGAAGGTGGTCATG 803
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1101 CTGTGGGGTACTGGATACATAAGGGAATGCCATCAGAGAAGGTGGTCATG 1150

804 GGCATCCCCACATATGGGCACTCCTTCACACTGGCCTCTGCAGAAACCAC 853
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1151 GGCATCCCCACATATGGGCACTCCTTCACACTGGCCTCTGCAGAAACCAC 1200

854 CGTGGGGGCCCCTGCCTCTGGCCCTGGAGCTGCTGGACCCATCACAGAGT 903
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1201 CGTGGGGGCCCCTGCCTCTGGCCCTGGAGCTGCTGGACCCATCACAGAGT 1250

904 CTTCAGGCTTCCTGGCCTATTATGAGATCTGCCAGTTCCTGAAAGGAGCC 953
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1251 CTTCAGGCTTCCTGGCCTATTATGAGATCTGCCAGTTCCTGAAAGGAGCC 1300

954 AAGATCACGAGGCTCCAGGATCAGCAGGTTCCCTACGCAGTCAAGGGGAA 1003
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1301 AAGATCACGAGGCTCCAGGATCAGCAGGTTCCCTACGCAGTCAAGGGGAA 1350
```

FIG.6B

```
1004 CCAGTGGGTGGGCTATGATGATGTGAAGAGTATGGAGACCAAGGTTCAGT 1053
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1351 CCAGTGGGTGGGCTATGATGATGTGAAGAGTATGGAGACCAAGGTTCAGT 1400

1054 TCTTAAAGAATTTAAACCTGGGAGGAGCCATGATCTGGTCTATTGACATG 1103
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1401 TCTTAAAGAATTTAAACCTGGGAGGAGCCATGATCTGGTCTATTGACATG 1450

1104 GATGACTTCACTGGCAAATCCTGCAACCAGGGCCCTTACCCTCTTGTCCA 1153
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1451 GATGACTTCACTGGCAAATCCTGCAACCAGGGCCCTTACCCTCTTGTCCA 1500

1154 AGCAGTCAAGAGAAGCCTTGGCTCCCTGTGAAGGATTAACTTACAGAGAA 1203
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1501 AGCAGTCAAGAGAAGCCTTGGCTCCCTGTGAaggattaacttacagagaa 1550

1204 GCAGGCAAGATGACCTTGCTGCCTGGGGCCTGCTCTCTCCCAGGAATTCT 1253
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1551 gcaggcaagatgaccttgctgcctggggcctgctctctcccaggaattct 1600

1254 CATGTGGGATTCCCCTTGCCAGGCCGGCCTTTGGATCTCTCTTCCAAGCC 1303
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1601 catgtgggattcccctgccaggccggcctttggatctctcttccaagcc 1650

1304 TTTCCTGACTTCCTCTTAGATCATAGATTGGACCTGGTTTTGTTTTCCTG 1353
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1651 tttcctgacttcctcttagatcatagattggacctggttttgttttcctg 1700

1354 CAGCTGATGACTTGTTGCCCTGAAGTACAATAAAAAAAATTCATTTTGCT 1403
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1701 cagctgatgacttgttgccctgaagtacaataaaaaaaattcattttgct 1750

1404 CCAAAAAAAAAAAAAAAAAAAACTCGAG..................... 1432
     ||||||||||||||||||||||||||||
1751 ccagtaaaaaaaaaaaaaaaaaactcgag..................... 1779
```

FIG.6C

```
  1 .....MDQKSLWAGVVVLLLLQG..........................  18
          ||||||||||||||||||
  1 MGATTMDQKSLWAGVVVLLLLQGEMGFCYVARAGLELLGSRSPPASASQS  50

 19 ....GSAYKLVCYFTNWSQDRQEPGKFTPENIDPFLCSHLIYSFASIENN  64
         ||||||||||||||||||||||||||||||||||||||||||||||||
 51 AGITGSAYKLVCYFTNWSQDRQEPGKFTPENIDPFLCSHLIYSFASIENN 100

 65 KVIIKDKSEVMLYQTINSLKTKNPKLKILLSIGGYLFGSKGFHPMVDSST 114
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
101 KVIIKDKSEVMLYQTINSLKTKNPKLKILLSIGGYLFGSKGFHPMVDSST 150

115 SRLEFINSIILFLRNHNFDGLDVSWIYPDQKENTHFTVLIHELAEAFQKD 164
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
151 SRLEFINSIILFLRNHNFDGLDVSWIYPDQKENTHFTVLIHELAEAFQKD 200

165 FTKSTKERLLLTAGVSAGRQMIDNSYQVEKLAKDLDFINLLSFDFHGSWE 214
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
201 FTKSTKERLLLTAGVSAGRQMIDNSYQVEKLAKDLDFINLLSFDFHGSWE 250

215 KPLITGHNSPLSKGWQDRGPSSYYNVEYAVGYWIHKGMPSEKVVMGIPTY 264
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
251 KPLITGHNSPLSKGWQDRGPSSYYNVEYAVGYWIHKGMPSEKVVMGIPTY 300

265 GHSFTLASAETTVGAPASGPGAAGPITESSGFLAYYEICQFLKGAKITRL 314
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
301 GHSFTLASAETTVGAPASGPGAAGPITESSGFLAYYEICQFLKGAKITRL 350

315 QDQQVPYAVKGNQWVGYDDVKSMETKVQFLKNLNLGGAMIWSIDMDDFTG 364
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
351 QDQQVPYAVKGNQWVGYDDVKSMETKVQFLKNLNLGGAMIWSIDMDDFTG 400

365 KSCNQGPYPLVQAVKRSLGSL. 385
    |||||||||||||||||||||
401 KSCNQGPYPLVQAVKRSLGSL* 422
```

FIG.6D

reference number 1488.046PC00 | TBA

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 5, line 16.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution

AMERICAN TYPE CULTURE COLLECTION

Address of depositary institution *(including postal code and country)*

12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
May 3, 1996	ATCC Designation 97529

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

DNA PLASMID 1290853 (Docket PF286)

In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer: Yvette E. Simms	Authorized officer

(DNA Plasmid, 1290853)

SINGAPORE

The applicant hereby request that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

NORWAY

The applicant hereby request that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the pbulic under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

INTERNATIONAL SEARCH REPORT

Inter ional application No.
PCT/US96/13003

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 15/56, 15/11, 15/10, 15/64, 1/21, 15/63
US CL :435/209, 320.1, 252.3, 172.3, 240.2; 536/23.1, 23.2
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/209, 320.1, 252.3, 172.3, 240.2; 536/23.1, 23.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, SCISEARCH, LIFESCI, BIOTECHDS, BIOSIS, EMBASE, HCAPLUS, NTIS, WPI search terms: chitinase#, gene# or sequence#, human

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	Database GenBank on STN, US National Institute of Health, (Bethesda MD, USA), GenBank Acc. No. N26985, HILLIER et al., 'The WashU-Merck EST Project', 29 December 1995, see sequence.	1, 8-14 --- 2-7, 15
X --- Y	Database GenBank on STN, US National Institute of Health, (Bethesda, MD, USA), GenBank Acc. No. N40107, HILLIER et al., 'The WashU-Merck EST Project', 22 January 1996, see sequence.	1, 8-14 --- 2-7, 15
X --- Y	Database GenBank on STN, US National Institute of Health, (Bethesda, MD, USA), GenBank Acc. No. T65854, HILLIER et al., 'WashU-Merck EST Project', 20 February 1995.	1, 8-14 --- 2-7, 15
X --- Y	Database GenBank on STN, US National Institute of Health, (Bethesda MD, USA), GenBank Acc. No. T66009, HILLIER et al., 'WashU-Merck EST Project', 20 February 1995.	1, 8-14 --- 2-7, 15

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 08 NOVEMBER 1996	Date of mailing of the international search report 22 NOV 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer REBECCA PROUTY Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

Inte.. .tional application No.
PCT/US96/13003

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	Database GenBank on STN, US National Institute of Health, (Bethesda MD, USA), GenBank Acc. No. F06990, AFFRAY et al., 'IMAGE: Integrated molecular analysis of the human genome and its expression', C.R. Acad. Sci. II, Sci. Vien., 318, 263-272, 15 February 1995, see sequence.	1, 8-14 --- 2-7, 15
X --- Y	Database GenBank on STN, US National Institute of Health, (Bethesda MD, USA), GenBank Acc. No. Z40100, AFFRAY et al., 'IMAGE: Integrated molecular analysis of the human genome and its expression', C.R. Acad. Sci. II, Sci. Vien., 318, 263-272, 04 November 1994, see sequence.	1, 8-14 --- 2-7, 15
X --- Y	Database GenBank on STN, US National Institute of Health, (Bethesda, MD, USA), GenBank Acc. No. N64805, HILLIER et al., 'The WashU-Merck EST Project', 01 March 1996, see sequence.	1, 8-14 --- 2-7, 15
X --- Y	Database GenBank on STN, US National Institute of Health, (Bethesda, MD, USA), GenBank Acc. No. H10989, HILLIER et al., 'The WashU-Merck EST Project', 26 June 1995, see sequence.	1, 8-14 --- 2-7, 15
Y	FUHRMAN et al. Transmission-Blocking Antibodies Recognize Microfilarial Chitinase in Brugian Lymphatic Filariasis. Proc. Natl. Acad. Sci. USA. 1992, Vol. 89, pages 1548-1552, see entire document.	2-7, 15
Y	HAKALA et al. Human Cartilage gp-39, a Major Secretory Product of Articular Chondrocytes and Synovial Cells, Is a Mammalian Member of a Chitinase protein Family. J. Biol. Chem. 05 December 1993, Vol. 268, No. 34, pages 25803-25810, see entire document.	2-7, 15

INTERNATIONAL SEARCH REPORT

Inte. ..tional application No.
PCT/US96/13003

Box I Observations where certain claims were found unsearchable (Continuation of item 1 f first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional ..earch fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-15

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Inte .tional application No.
PCT/US96/13003

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claims 1-15, drawn to DNA, vectors, host cells and expression of human chitinase alpha and alpha-2.

Group II, claims 16, 17 and 19, drawn to human chitinase alpha and alpha-2 protein and method of use.

Group III, claim 18, drawn to human chitinase alpha and alpha-2 antibody.

Group IV, claim 20, drawn to a method of treating a tissue remodelling disorder with a human chitinase alpha or alpha-2 antagonist.

The inventions listed as Groups I-IV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: each of the four groups recites a chemically unrelated special technical feature. The special technical feature of Group I is a nucleic acid encoding human chitinase alpha or alpha-2, the special technical feature of Group II is a human chitinase alpha or alpha-2 protein, the special technical feature of Group III is a human chitinase alpha or alpha-2 antibody and the special technical feature of Group IV is a human chitinase alpha or alpha-2 antagonist.